| ERIC Document Resume | See Instruction Sheet | DATE OF RESUME | DATE RECEIVED |
|---|---|---|---|

A. PUBLICATION DATA

1. ERIC NUMBER: ED001177
2. ORIGINATOR CONTROL NO(S).: N.A.
3. IDENTIFICATION NO.: CR 577

DO NOT WRITE IN SPACE BELOW

MICROFILM COPY AVAILABLE: YES ☒ NO ☐

COPYRIGHT GRANTED: YES ☐ NO ☒

COPYRIGHT RELEASE FILED: YES ☐ NO ☐

TYPE OF REPORT:
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☒
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

4. AUTHOR(S): Borg, Walter R.
5. TITLE OF DOCUMENT ABSTRACTED: An Evaluation of Ability Grouping
6. PUBLICATION TITLE: N.A.
7. EDITOR(S): N.A.
8. PLACE OF ORIGIN
ORG. Utah State Univ.
Logan, Utah
ST.
CITY, STATE
TELE. EXT.
INST/LOC CODE
9. PUBLISHER OR SOURCE: Cooperative Research Program
Office of Edu. - Dept. of HEW - Wash.

| 10. VOL. | 11. PAGES | 12. MO./YR. | 13. TOTAL PP. |
|---|---|---|---|
| N.A. | N.A. | 64 | 462 |

B. INDEXING DATA

14. SUBJECT AREA
15. TYPE OF ENDEAVOR: research

16. INDEXING TERMS (6-12)

| | | | | |
|---|---|---|---|---|
| | | aspiration | | |
| study habits | attitude | personality | | |
| achievement | values | problems | | |
| potential | self-concept | ability grouping | enrichment | |

17. ABSTRACT (MAX. 250 WORDS)

The project's sample consisted of children of elementary, junior high, and high school age. They were divided into two groups: 1) an ability grouping that differentiates the curriculum principally by adjusting the rate of presentation of curricular material, 2) random grouping that differentiates the curriculum principally through the use of enrichment.

The two groups were examined and evaluated for: 1) Study Habits, 2) Achievement, 3) Realization of Achievement Potential, 4) Pupil Attitutudes, 5) Pupil Problems, 6) Self-Concept, 7) Aspiration level and achievement values, 8) Personality.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION
20. DIVISION OR DEPT.
21. CONTRACT NO.
PROJECT NO.
22. TOTAL COST OF PROJECT
23. DURATION OR PROJECT

U.S. Department of Health, Education, and Welfare
Office of Education

7

ED001177

C R P
No. 577

BORG

UTAH
STATE
UNIV.

# AN EVALUATION OF ABILITY GROUPING

## COOPERATIVE RESEARCH PROJECT NO. 577

WALTER R. BORG

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

UTAH STATE UNIVERSITY
LOGAN, UTAH

# AN EVALUATION OF ABILITY GROUPING

Cooperative Research Project No. 577

by

Walter R. Borg
Utah State University
Logan, Utah

1964

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education and Welfare.

# ACKNOWLEDGMENT

To carry out a field study of the magnitude reported in this volume requires the interest, cooperation and active help of a great many individuals. It would be impossible for me to acknowledge everyone who has assisted in some way in the planning and execution of this research. Certain persons, however, played such important roles in the project that they must be singled out from among the great number who participated. First among those whose support was essential to the success of this undertaking are Dr. T. O. Smith and Dr. T. H. Bell, the superintendents in the two school districts that cooperated in the project. These men not only helped in all of the usual ways that the chief administrator in a school district can help a researcher, but consistently contributed more than could be expected toward adding to the objectivity and validity of the study.

The three men assigned to provide liaison between the school districts and the investigators were Dr. Lionel Drexel, Mr. Vern Call, and Mr. J. Rhees. They were invariably ready to work with the investigators in overcoming the many problems that arose during the more than four years over which data were collected. Without the willing help of these three capable educators, it would have been impossible to successfully complete the project.

In her capacity as general research assistant, Miss Mary Ann Sharp supervised the work of the clerical staff, and assumed much of the responsibility for preparing test schedules, assembling materials, and keeping records. Although an undergraduate during much of the time that the research was underway, she carried out her many duties in such a way as to earn the admiration and gratitude of all persons involved in the project.

Finally, much credit must go to the graduate research assistants who carried out various parts of the Utah study in conjunction with their graduate training. These include Dr. Luna Brite, Dr. David Brinks, Dr. George Jeffs, Dr. Nephi Manning, Dr. Clarence Swaner, Mr. Val Christensen, Mr. Luan H. Ferrin, Mr. John N. Giboney, and Mr. Benjamin Standing.

--Walter R. Borg
Project Director

## TABLE OF CONTENTS

Chapter
No. Page

1. INTRODUCTION TO THE UTAH STUDY . . . 1
    The Problem . . . 1
    Research Procedures . . . 2
    Analysis . . . 3
    Organization of the Report . . . 4

2. THE SAMPLE . . . 7
    Introduction . . . 7
    District Comparisons . . . 8
        Financial Resources . . . 8
        Comparability of Teachers in the Two Districts . . . 8
        Similarities and Differences in Curriculum . . . 10
        Requirements at Secondary Level . . . 11
            Junior High School Mathematics . . . 11
            Junior High School Science . . . 11
            High School Mathematics . . . 12
            High School Science Requirements . . . 12
            Summary . . . 12
        Selecting Comparable Schools . . . 13
    Selecting Research Subjects . . . 13
            Sample IV . . . 15
            Sample VI . . . 16
            Sample VII . . . 16
            Sample VIII . . . 19
            Sample IX . . . 19
            Sample Variations . . . 19
    Comparability of Research Subjects . . . 20
            Socioeconomic Class Differences Between District A and District R Pupils of Comparable Ability . . . 20
            Intelligence Differences at Beginning of the Study . . . 22

3. THE TREATMENTS . . . 25
    Introduction . . . 25
    The Ability Grouping System . . . 25
            Elementary . . . 25
            Secondary . . . 26
    The Enrichment System . . . 27
    Reduction in Heterogeneity in District A . . . 27
        Introduction . . . 27
        Reduction Achieved . . . 28
    Curriculum Adjustments . . . 29
        Introduction . . . 29
        Adapting the Curriculum . . . 35
        Curriculum Adjustments in District A Elementary Classes . . . 36

        Reading . . . . . . . . . . . . . . . . . . 36
        Arithmetic . . . . . . . . . . . . . . . . . 36
        Science . . . . . . . . . . . . . . . . . . 37
        Social Studies . . . . . . . . . . . . . . . 37
    Curriculum Adjustments in District R Elementary Classes . . 37
        Reading . . . . . . . . . . . . . . . . . . 37
        Arithmetic . . . . . . . . . . . . . . . . . 38
        Science . . . . . . . . . . . . . . . . . . 38
        Social Studies . . . . . . . . . . . . . . . 38
  Curriculum Adjustments in the Junior High Schools . . . . . . 38
    District A Adjustments . . . . . . . . . . . . . . 38
        Mathematics . . . . . . . . . . . . . . . . 38
        Science . . . . . . . . . . . . . . . . . . 39
    District R Adjustments . . . . . . . . . . . . . . 39
        Mathematics . . . . . . . . . . . . . . . . 39
        Science . . . . . . . . . . . . . . . . . . 40
  Curriculum Adjustments in the High Schools . . . . . . . . 40
    Limitations at the Secondary Level . . . . . . . . . 41
  Summary of Sampling and Treatment Differences . . . . . . . 41

4. ACHIEVEMENT . . . . . . . . . . . . . . . . . . . . 43
  Introduction . . . . . . . . . . . . . . . . . . . 43
  Working Hypotheses . . . . . . . . . . . . . . . . . 43
  Previous Research Related to Achievement and Grouping . . . . . 44
        Introduction . . . . . . . . . . . . . . . . 44
    Research at the Elementary Level . . . . . . . . . . 45
    Research at the Secondary Level . . . . . . . . . . 53
    Critical Evaluation . . . . . . . . . . . . . . . 60
  Procedure . . . . . . . . . . . . . . . . . . . . 62
    Sample . . . . . . . . . . . . . . . . . . . . 62
    Achievement Measures . . . . . . . . . . . . . . 62
  Achievement Results . . . . . . . . . . . . . . . . 63
    Original Fourth Grade Sample (Sample IV) . . . . . . . . 63
        Introduction . . . . . . . . . . . . . . . . 63
        First Year Achievement . . . . . . . . . . . . 64
        Achievement in Specific Areas . . . . . . . . . 66
        A Note of Caution . . . . . . . . . . . . . . 66
        Second Year Achievement . . . . . . . . . . . 67
        Third Year Achievement . . . . . . . . . . . . 70
        Fourth Year Achievement . . . . . . . . . . . 70
        Overall Results for the Four Years . . . . . . . . 71
        Results for the Last Three Years . . . . . . . . 71
        Conclusions . . . . . . . . . . . . . . . . . 76
    Original Sixth Grade Sample (Sample VI) . . . . . . . . 76
        First Year Achievement . . . . . . . . . . . . 76
        Specific Achievement Areas . . . . . . . . . . 77
        Second Year's Achievement . . . . . . . . . . 77
        Third and Fourth Year's Achievement . . . . . . . . 80
        Overall Results for the Four Years . . . . . . . . 80

Original Seventh Grade Sample (Sample VII) . . . . . . . 80
First Year's Achievement . . . . . . . . . . 80
Second Year's Achievement . . . . . . . . . . 85
Third and Fourth Year's Achievement . . . . . . . 85
Overall Results for the Four Years . . . . . . . . 85
Original Eighth Grade Sample (Sample VIII) . . . . . . . 89
First Year's Achievement . . . . . . . . . . 89
Second Year's Achievement . . . . . . . . . . 89
Third and Fourth Year's Achievement . . . . . . . 89
Overall Results for the Four Years . . . . . . . . 91
Original Ninth Grade Sample (Sample IX) . . . . . . . 91
First Year's Achievement . . . . . . . . . . 91
Second Year's Achievement . . . . . . . . . . 91
Third and Fourth Year's Achievement . . . . . . . 95
Overall Results for the Four Years . . . . . . . . 95
Achievement Results for a Sample of Superior Junior High School Students . . . . . . . . . . . . . . . . . . . 95
Introduction . . . . . . . . . . . . . . . 95
Procedure . . . . . . . . . . . . . . . . . 98
Results . . . . . . . . . . . . . . . . . . 98
Mathematics Achievement . . . . . . . . . . 98
Science Results . . . . . . . . . . . . . 100
Discussion . . . . . . . . . . . . . . . 100
Summary and Conclusions . . . . . . . . . . . . . 101
Introduction . . . . . . . . . . . . . . . 101
Elementary School Achievement . . . . . . . . . 101
Junior High School Achievement . . . . . . . . . 104
Senior High School Achievement . . . . . . . . . 104
Achievement Summary by Ability Level . . . . . . . 105

5. REALIZATION OF ACHIEVEMENT POTENTIAL . . . . . . . . . . 109
Introduction . . . . . . . . . . . . . . . . . . 109
Working Hypotheses . . . . . . . . . . . . . . . . 110
Previous Research Related to Underachievement . . . . . . . 110
Introduction . . . . . . . . . . . . . . . . 110
Environmental Factors . . . . . . . . . . . . . 111
Personal Factors . . . . . . . . . . . . . . . 112
Summary . . . . . . . . . . . . . . . . . . 115
Procedure . . . . . . . . . . . . . . . . . . . 116
Sample . . . . . . . . . . . . . . . . . . 116
Measures Used . . . . . . . . . . . . . . . . 118
California Short-Form Test of Mental Maturity . . . . 118
California Achievement Test Battery (CAT) . . . . . 118
California Study Methods Survey (CSMA) . . . . . 120
Thinking About Yourself . . . . . . . . . . . 120
Thematic Apperception Test (TAT) . . . . . . . . 121
Other Measures . . . . . . . . . . . . . . . . 121
Social Class . . . . . . . . . . . . . . . 121
Health . . . . . . . . . . . . . . . . . . 121

Results . . . 122
Introduction . . . 122
Sex Differences . . . 122
Grade Differences . . . 123
Between-treatment Differences . . . 123
Comparisons Between Normal Achievers and Underachievers 125
Study Methods . . . 127
Emotional Disturbance . . . 129
Achievement Motivation . . . 131
Socioeconomic Status . . . 131
Parents' Attitude Towards School . . . 133
Health . . . 133
Summary and Conclusions . . . 134
First Phase . . . 134
Second Phase . . . 136

6. STUDY HABITS . . . 141
Introduction . . . 141
Working Hypotheses . . . 141
Procedure . . . 142
Sample . . . 142
Measures . . . 142
California Study Methods Survey (CSMS) . . . 142
Survey of Study Habits and Attitudes . . . 142
Results . . . 142
Sample IV Data . . . 142
Introduction . . . 142
Limitations in the Sample IV Data . . . 144
Results . . . 144
Summary . . . 146
Discussion . . . 146
Sample VI Data . . . 147
Introduction . . . 147
Results . . . 147
Discussion . . . 149
Sample VIII Data . . . 150
Introduction . . . 150
Results . . . 150
Discussion . . . 152
Sample IX Data . . . 152
Introduction . . . 152
Results . . . 152
Discussion . . . 153
Ability Level Differences in Study Habit Scores . . . 155
Introduction . . . 155
Results . . . 155
Summary and Conclusions . . . 157
Elementary . . . 157
Secondary . . . 157
Level Differences . . . 157

7. SOCIOMETRIC CHOICE . . . 159
- Introduction . . . 159
- Working Hypotheses . . . 161
    - Sociometric Status Patterns . . . 162
    - Accuracy of Perceived Status . . . 162
    - Mobility of Sociometric Status . . . 162
    - Accuracy of Perceived Status over a Three Year Period . . . 163
    - Relationships Between Sociometric Variables and Other Measures . . . 163
- Previous Research Related to Grouping and Sociometric Choice . . . 163
    - Sociometry . . . 163
    - Sociometric Terms . . . 164
    - Ability and Sociometric Status . . . 164
    - Social Aspects of Ability Grouping . . . 168
- Procedure . . . 172
    - Sample . . . 172
    - Measures Used . . . 172
    - Method of Analysis . . . 174
- Results . . . 175
    - Sociometric Choice Patterns . . . 175
        - Between-treatment Differences . . . 175
        - Level Differences . . . 178
        - Sex Differences . . . 179
    - Accuracy of Perceived Status . . . 180
        - Between-treatment Differences . . . 180
        - Level Differences . . . 180
        - Sex Differences . . . 180
    - Mobility in Sociometric Status . . . 183
        - Between-treatment Differences . . . 183
        - Level Differences . . . 187
        - Sex Differences . . . 187
        - Long Range Stability of the Accuracy of Perceived Status Score . . . 187
        - Relationships Between Sociometric Status and Other Variables . . . 189
- Summary and Conclusions . . . 192

8. PUPIL ATTITUDES . . . 199
- Introduction . . . 199
- Working Hypotheses . . . 199
    - Attitude Towards Peers . . . 200
    - Attitude Towards the Teacher . . . 200
    - Attitude Towards School . . . 201
- Previous Research Related to Pupil Attitudes and Grouping . . . 202
- Procedure . . . 204
    - Sample . . . 204
    - Measures . . . 204
        - The USU School Inventory . . . 204

Attitude Towards Peers . . . . . . . . . . . . . 206
Attitude Towards the Teacher . . . . . . . . . 206
Attitude Towards School . . . . . . . . . . . . 210
School Problems . . . . . . . . . . . . . . . . 212
School Relations . . . . . . . . . . . . . . . . 212
Summary and Conclusions . . . . . . . . . . . . . 214

9. PUPIL PROBLEMS . . . . . . . . . . . . . . . . . 217
Introduction . . . . . . . . . . . . . . . . . . 217
Working Hypotheses . . . . . . . . . . . . . . . 217
Procedure . . . . . . . . . . . . . . . . . . . 218
Subjects . . . . . . . . . . . . . . . . . . 218
Measures . . . . . . . . . . . . . . . . . . 219
Results . . . . . . . . . . . . . . . . . . . . 219
Looking Ahead . . . . . . . . . . . . . . . 219
Getting Along with Others . . . . . . . . . 222
Boy Meets Girl . . . . . . . . . . . . . . . 226
Home and Family . . . . . . . . . . . . . . 228
Health . . . . . . . . . . . . . . . . . . . 228
Things in General . . . . . . . . . . . . . 229
Serious Problems and Basic Difficulties . . . . . 229
Summary and Conclusions . . . . . . . . . . . . . 230

10. SELF-CONCEPT . . . . . . . . . . . . . . . . . . 233
Introduction . . . . . . . . . . . . . . . . . . 233
Working Hypotheses . . . . . . . . . . . . . . . 234
Between-Treatment Differences . . . . . . . . . 235
Within-Treatment Differences . . . . . . . . . 235
Previous Research Related to Self-Concept and Grouping . . . 236
Self-Concept and Ability Grouping . . . . . . . 236
Other Pertinent Research . . . . . . . . . . . 240
Procedure . . . . . . . . . . . . . . . . . . . 244
Sample and Measures . . . . . . . . . . . . . 244
The Index of Adjustment and Values . . . . . . 244
Intercorrelational Data . . . . . . . . . . . 245
Reliability . . . . . . . . . . . . . . . . . 245
Sense of Personal Worth . . . . . . . . . . . 247
Self-Acceptance . . . . . . . . . . . . . . . 247
About Myself Scale . . . . . . . . . . . . . 247
Results . . . . . . . . . . . . . . . . . . . . 249
Sample IV . . . . . . . . . . . . . . . . . . 249
Introduction . . . . . . . . . . . . . . . . 249
Sex Differences on Elementary IAV . . . . . . 249
Classification by Ability . . . . . . . . . . 249
Sex Differences on Junior High School IAV . . . . 250
Concept of Self Data . . . . . . . . . . . . 250
Discussion . . . . . . . . . . . . . . . . . 254
Acceptance of Self Data . . . . . . . . . . . 255
Discrepancy Score Data . . . . . . . . . . . 257
Ideal Self Data . . . . . . . . . . . . . . . 261

Self Concept Changes Accompanying Changes in Ability Level in District A . . . . . . . . . . . . . . . . 261
Concept of Self . . . . . . . . . . . . . . . 264
Acceptance of Self . . . . . . . . . . . . . . 264
Ideal Self . . . . . . . . . . . . . . . . 264
Discrepancy Score . . . . . . . . . . . . . . 264
Sample VI Results . . . . . . . . . . . . . . . . 265
Introduction . . . . . . . . . . . . . . . 265
Sex Differences . . . . . . . . . . . . . . 265
Concept of Self Data . . . . . . . . . . . . . 267
Acceptance of Self Data . . . . . . . . . . . . 267
Discrepancy Score Data . . . . . . . . . . . . 270
Ideal Self Data . . . . . . . . . . . . . . 270
Self-Concept Changes Accompanying Ability Level Changes in District A . . . . . . . . . . . . . . . . 273
Introduction . . . . . . . . . . . . . . . 273
Concept of Self . . . . . . . . . . . . . . . 273
Acceptance of Self . . . . . . . . . . . . . . 273
Ideal Self . . . . . . . . . . . . . . . . 275
Discrepancy Score . . . . . . . . . . . . . . 275
Sample VII Results . . . . . . . . . . . . . . . . 275
Sex Differences . . . . . . . . . . . . . . 275
Concept of Self Data . . . . . . . . . . . . . 278
Acceptance of Self Data . . . . . . . . . . . . 283
Discrepancy Score Data . . . . . . . . . . . . 283
Ideal Self Data . . . . . . . . . . . . . . 283
Sample VIII Results . . . . . . . . . . . . . . . . 284
Sex Differences . . . . . . . . . . . . . . 284
Concept of Self Data . . . . . . . . . . . . . 284
Acceptance of Self Data . . . . . . . . . . . . 284
Discrepancy Score Data . . . . . . . . . . . . 290
Ideal Self Data . . . . . . . . . . . . . . 290
Sample IX Results . . . . . . . . . . . . . . . . 290
Introduction . . . . . . . . . . . . . . . 290
Sex Differences . . . . . . . . . . . . . . 290
Concept of Self Data . . . . . . . . . . . . . 291
Acceptance of Self Data . . . . . . . . . . . . 291
Discrepancy Score Data . . . . . . . . . . . . 291
Ideal Self Data . . . . . . . . . . . . . . 291
Discussion . . . . . . . . . . . . . . . . 291
Results on Other Self-Concept Measures . . . . . . . . . 297
Sense of Personal Worth . . . . . . . . . . . 297
Self-Acceptance . . . . . . . . . . . . . . 297
About Myself Scale . . . . . . . . . . . . . 298
Summary and Conclusions . . . . . . . . . . . . . . 300
Between-treatment Differences . . . . . . . . . . . 300
Concept of Self . . . . . . . . . . . . . . . 300
Acceptance of Self . . . . . . . . . . . . . . 302
Discrepancy Score . . . . . . . . . . . . . . 302
Ideal Self . . . . . . . . . . . . . . . . 303

Within-treatment Differences . . . . . . . . . . . 303
Self-Concept Changes Accompanying Ability Level Changes in District A . . . . . . . . . . 308

11. ASPIRATION LEVEL AND ACHIEVEMENT VALUES . . . . . . 311
Introduction . . . . . . . . . . . . . . . 311
Working Hypotheses . . . . . . . . . . . . . 312
Between-treatment Differences . . . . . . . 312
Level Differences . . . . . . . . . . . 312
Previous Research Pertinent to Aspiration Level and Ability Grouping . . . . . . . . . . . . . . . . 314
Introduction . . . . . . . . . . . . 314
Group Influences on Level of Aspiration . . . . 314
Level of Aspiration and Achievement . . . . . 316
Level of Aspiration Changes Following Success and Failure . . . . . . . . . . . . . 318
Procedure . . . . . . . . . . . . . . . . . 318
Subjects . . . . . . . . . . . . . . . 318
Introduction . . . . . . . . . . . . 318
District A Sample . . . . . . . . . . . 319
District R Sample . . . . . . . . . . . 319
Measures Employed . . . . . . . . . . . . 321
Introduction . . . . . . . . . . . . 321
Classroom Aspiration . . . . . . . . . . 321
Social Aspiration . . . . . . . . . . . 321
Educational Aspiration . . . . . . . . . 322
Occupational Aspiration . . . . . . . . . 322
Value Achievement . . . . . . . . . . 322
Laboratory Aspiration . . . . . . . . . . 323
Results . . . . . . . . . . . . . . . . . 323
Introduction . . . . . . . . . . . . 323
Between-Treatment Differences (Sample IV) . . 323
Between-Treatment Differences (Sample VI) . . 326
Level Differences . . . . . . . . . . . 326
Summary and Conclusions . . . . . . . . . . . 329

12. PERSONALITY . . . . . . . . . . . . . . . . 333
Introduction . . . . . . . . . . . . . . . 333
Working Hypotheses . . . . . . . . . . . . . 334
California Test of Personality . . . . . . . 334
Inventory of Factors GAMIN . . . . . . . . 335
Objective Analytic Personality Factors . . . . 335
California Psychological Inventory . . . . . . 336
Previous Research Related to Personality and Grouping . . 337
Introduction . . . . . . . . . . . . 337
Ability Grouping . . . . . . . . . . . 339
Procedure . . . . . . . . . . . . . . . . . 342
Personality Measures . . . . . . . . . . . 342

California Test of Personality (CTP) 342
Inventory of Factors GAMIN 342
Objective-Analytic Personality Test Factors 344
The California Psychological Inventory 345
Sample 346
Results 346
California Test of Personality 346
Withdrawing Tendencies 349
Anti-Social Tendencies 349
Total Adjustment 349
Inventory of Factors GAMIN 356
Ascendancy 356
Inferiority 360
Objective-Analytic Personality Test Factors 360
Competent Assertiveness 360
Anxiety to Achieve 362
California Psychological Inventory (CPI) 363
Class I Variables 363
Class II Variables 372
Class III Variables 377
Class IV Variables 377
Summary and Conclusions 380
California Test of Personality (CTP) 380
Inventory of Factors GAMIN 387
Objective-Analytic Personality Test Factors (OA) 387
California Psychological Inventory (CPI) 388

13. PERSONALITY - THE PROJECTIVE STUDIES 391
Introduction 391
Working Hypotheses 391
Procedure 393
The Measures 393
Introduction 393
The Individual Projective Test 394
The Group Projective Test 394
Scoring the Projective Measure 396
Anxiety Measurement in the Sixth Grade Study 397
The Children's Manifest Anxiety Scale (CMAS) 398
The Test Anxiety Scale for Children (TASC) 399
The Sample 400
Fifth Grade Study 400
Sixth Grade Study 400
The Pilot Studies 402
Fifth Grade 402
Sixth Grade 403
Results 404
Fifth Grade Study 404
Sixth Grade Study 407
Projective Data 407
Anxiety Data 409

Ability Level Comparisons . . . . . . . . . . . 412
Sex Differences . . . . . . . . . . . . . . . 415
Summary and Conclusions . . . . . . . . . . . . . 417

APPENDICES . . . . . . . . . . . . . . . . . . . . . . . . 421

# LIST OF TABLES

Table No. | | Page
2-1. Instructional Budgets Per ADA . . . . . . . . . . . . . 9

2-2. Number of Pupils in Each Sample . . . . . . . . . . . 14

2-3. Distribution of District A Elementary Sample, 4th Grade on CAT First Battery . . . . . . . . . . . . . . . . 17

2-4. Distribution of District A Elementary Sample, 6th Grade on CAT First Battery . . . . . . . . . . . . . . . . 18

2-5. Socioeconomic Class Differences Between District A and District R Pupils of Comparable Ability . . . . . . . . . . 21

2-6. CMM Language and Non-language Intelligence Grade Placements . 23

3-1. Distribution of District R Elementary Sample, 4th Grade on CAT First Battery . . . . . . . . . . . . . . . . 30

3-2. Distribution of District R Elementary Sample, 6th Grade on CAT First Battery . . . . . . . . . . . . . . . . 32

3-3. California Achievement Test -- First Battery Range of SD's and Average SD's for Classrooms in the Two Treatments . . . . 34

4-1. Ability Grouping Studies Concerned with Achievement at the Elementary School Level . . . . . . . . . . . . . 46

4-2. Ability Grouping Studies Concerned with Achievement at the Secondary School Level . . . . . . . . . . . . . 54

4-3. Achievement of Sample IV Pupils During Their Fourth Grade Year . 65

4-4. Achievement of Sample IV Pupils During Their Fifth Grade Year . 68

4-5. Achievement of Sample IV Pupils During Their Sixth Grade Year . 72

4-6. Achievement of Sample IV Pupils During Their Seventh Grade Year . 74

4-7. Achievement of Sample VI Pupils During Their Sixth Grade Year . 78

4-8. Achievement of Sample VI Pupils During Their Seventh Grade Year 81

4-9. Achievement of Sample VI Pupils During Their Eighth Ninth Grade Years . . . . . . . . . . . . . . . . 82

4-10. Achievement of Sample VI Pupils Over a Four-Year Period . . . 83

4-11. Achievement of Sample VII Pupils During Their Seventh Grade Year 84

4-12. Achievement of Sample VII Pupils During Their Eighth Grade Year . 86

4-13. Achievement of Sample VII Pupils During Their Ninth and Tenth Grade Years . . . . . . . . . . . . . . . . 87

4-14. Achievement of Sample VII Pupils Over a Four-Year Period . . . 88

4-15. Achievement of Sample VIII Pupils During Their Eighth and Ninth Grade Years . . . . . . . . . . . . . . 90

4-16. Achievement of Sample VIII Pupils During Their Tenth and Eleventh Grade Years . . . . . . . . . . . . . 92

4-17. Achievement of Sample VIII Pupils Over a Four-Year Period . . 93

4-18. Achievement of Sample IX Pupils During Their Ninth and Tenth Grade Years . . . . . . . . . . . . . . . . 94

4-19. Achievement of Sample IX Pupils During Their Eleventh and Twelfth Grade Years . . . . . . . . . . . . . 96

4-20. Achievement of Sample IX Pupils Over a Four-Year Period . . . 97

4-21. Mathematics and Science Achievement of Superior Junior High School Pupils in Ability Grouped and Random Grouped Classes . 99

4-22. Summary of Achievement Differences Between District A and District R Samples . . . . . . . . . . . . . . 102

5-1. Percentage of Each Subgroup Classified as Overachievers, Normal Achievers and Underachievers . . . . . . . . . . . 117

5-2. Samples of Underachievers and Normal Achievers used in the 1959-60 Phase of the Study. . . . . . . . . . . . 119

5-3. Differences Between Boys and Girls in the Distribution of Overachievers, Normal Achievers and Underachievers . . . . . 124

5-4. Differences Between District A and District R Pupils in the Distribution of Overachievers, Normal Achievers and Underachievers 126

5-5. Tabulation of Health Data . . . . . . . . . . . . 135

6-1. Samples Employed in the Study Habits Phase of the Utah Study . . 143

6-2. Mean Scores of Sample IV Subgroups on the California Study Methods Survey . . . . . . . . . . . . . . . . . . 145

6-3. Mean Scores of Sample VI Subgroups on the California Study Methods Survey . . . . . . . . . . . . . . . . . . 148

6-4. Mean SSHA Scores of Sample VIII Subgroups . . . . . . . . 151

6-5. Mean SSHA Scores of Sample IX Subgroups . . . . . . . . . 154

6-6. Differences in Study Habits Scores Between Pupils of Different Ability Levels in Each District . . . . . . . . . . . . 156

7-1. Correlation Coefficients Between Group Social Acceptance and Mental Ability . . . . . . . . . . . . . . . . . . 167

7-2. Per Cent of Students in Achievement Groups Receiving Intellectual and Social Nominations in Homogeneous and Heterogeneous Classes 170

7-3. Numbers of Subjects in Each Subgroup for Each Replication of Sociometric Study . . . . . . . . . . . . . . . . . . 173

7-4. Differences Between Different SubGroups in the Four Replications Based Upon Chi-square Contingency Tables . . . . . . . . 176

7-5. Subgroup Mean Scores for Accuracy of Perceived Status . . . . 181

7-6. Between-treatment Differences in Accuracy of Perceived Status . . 182

7-7. Significant Differences Between Mean Scores on Accuracy of Perceived Status for Pupils of the Different Levels of Ability . . 184

7-8. Between-treatment Differences in Sociometric Status Mobility at Each Ability Level . . . . . . . . . . . . . . . . . . 185

7-9. Mobility and Stability of Sociometric Status . . . . . . . . 188

7-10. Correlation Coefficients of Accuracy of Perceived Status Scores Obtained in Different Years . . . . . . . . . . . . . 190

7-11. Test-Retest Correlation Coefficients of First Criterion Accuracy of Perceived Status Scores . . . . . . . . . . . . . 191

7-12. Correlations Between Sociometric Status and Other Variables for Each District . . . . . . . . . . . . . . . . . . 193

8-1. Sample Items From the USU School Inventory . . . . . . . . 205

8-2. Between-treatment Differences on the USU School Inventory . . 207

8-3. Within-treatment Differences on the USU School Inventory . . . 208

8-4. Sex Differences on the USU School Inventory . . . . . . . . 211

8-5. Between-treatment Differences on the SRA My School Scale . . . 213

9-1. Between-treatment Differences on SRA Junior and Youth Inventories 220

9-2. Differences Between Pupils at Different Ability Levels on SRA Junior and Youth Inventories . . . . . . . . . . . . 223

9-3. Sex Differences in Mean Problem Scores on the SRA Junior and Youth Inventories . . . . . . . . . . . . . . . . 225

10-1. Comparison of Mean Ability Self-Concept Ratings by Achievement Group and Sex in Homogeneous and Heterogeneous Classes . . 238

10-2. Comparison of Ratings of Concept-of-Self-as-School-Learner by Achievement Group and Sex and Between Homogeneous and Heterogeneous Classes . . . . . . . . . . . . . . 239

10-3. Comparison of Ratings of Concept-of-Self-as-School-Learner Between Achievement Groups in Homogeneous and Heterogeneous Classes . . . . . . . . . . . . . . . . . 241

10-4. Intercorrelational Matrices for IAV Scores . . . . . . . . . 246

10-5. Correlations Between IAV Scores Administered in 5th, 6th, and 7th Grades to Sample IV, District R (N = 455) . . . . . . . . 248

10-6. Sex Differences on Junior High School IAV Scores for Sample IV . 251

10-7. Sample IV - Differences in Concept of Self Scores Over a Three Year Period . . . . . . . . . . . . . . . . . 253

10-8. Sample IV - Differences in Acceptance of Self over a Three Year Period . . . . . . . . . . . . . . . . . . 256

10-9. Sample IV - Differences in Discrepancy Scores over a Three Year Period . . . . . . . . . . . . . . . . . . 258

10-10. Sample IV - Zero and Negative Discrepancy Scores by Ability Level and Treatment . . . . . . . . . . . . . . . 260

10-11. Sample IV - Differences in Ideal Self Scores over a Three Year Period . . . . . . . . . . . . . . . . . . 262

10-12. Sample IV - Changes in Ability Level, 5th to 6th Grade in District A Sample, and Accompanying Changes in IAV Scores . . . . . 263

10-13. Sex Differences on Junior High School IAV Scores for Sample VI . . 266

10-14. Sample VI - Differences in Concept of Self Scores over a Three Year Period . . . . . . . . . . . . . . . . 268

10-15. Sample VI - Differences in Acceptance of Self Scores over a Three Year Period . . . . . . . . . . . . . . . 269

10-16. Sample VI - Differences in Discrepancy Scores over a Three Year Period . . . . . . . . . . . . . . . . . . 271

10-17. Sample VI - Differences in Ideal Self Scores over a Three Year Period . . . . . . . . . . . . . . . . . . 272

10-18. Sample VI - Changes in Ability Level, 6th to 7th Grade in the District A Sample, and Accompanying Changes in IAV Scores . . 274

10-19. Sex Differences in Junior High School IAV Scores for Sample VII . 277

10-20. Sample VII - Differences in Concept of Self Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 279

10-21. Sample VII - Differences in Acceptance of Self Scores over a Two Year Period . . . . . . . . . . . . . . . . 280

10-22. Sample VII - Differences in Discrepancy Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 281

10-23. Sample VII - Differences in Ideal Self Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 282

10-24. Sex Differences in the IAV Scores for Sample VIII . . . . . . . 285

10-25. Sample VIII - Differences in Concept of Self Scores over a Two Year Period . . . . . . . . . . . . . . . . 286

10-26. Sample VIII - Differences in Acceptance of Self Scores over a Two Year Period . . . . . . . . . . . . . . . . 287

10-27. Sample VIII - Differences in Discrepancy Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 288

10-28. Sample VIII - Differences in Ideal Self Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 289

10-29. Sex Differences for Sample IX . . . . . . . . . . . . . 292

10-30. Sample IX - Differences in Concept of Self Scores over a Two Year Period . . . . . . . . . . . . . . . . . 293

10-31. Sample IX - Differences in Acceptance of Self Scores over a Two Year Period . . . . . . . . . . . . . . . . . 294

10-32. Sample IX - Differences in Discrepancy Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 295

10-33. Sample IX - Differences in Ideal Self Scores over a Two Year Period . . . . . . . . . . . . . . . . . . 296

12-34. Between-treatment Differences on the SRA About Myself Scale . . 299

10-35. Summary of Significant Between-treatment Differences on the IAV . 301

10-36. Summary of Significant Within Treatment Differences on the IAV . . 304

10-37. Summary of Significant Within Treatment Differences on the IAV . . 306

11-1. Number of Subjects at each Level in District A Grades Seven and Nine . . . . . . . . . . . . . . . . . . 320

11-2. Aspiration Level Means for Sample IV, During Their 7th Grade Year . 324

11-3. Directional Trends in Aspiration Level Measures for Sample IV . . 325

11-4. Aspiration Level Means for Sample VI, During Their 9th Grade Year . 327

11-5. Directional Trends in Aspiration Level Measures for Sample VI . . 328

12-1. Personality Variables Measured on the CPI . . . . . . . . 338

12-2. Reliability Coefficients of the California Test of Personality . . . 343

12-3. Test-Retest Reliability Coefficients on the CPI . . . . . . . 347

12-4. Samples Administered Personality and Related Measures in the Utah Study . . . . . . . . . . . . . . . . . 348

12-5. Sample IV: Between-treatment Differences on the California Test of Personality Variables . . . . . . . . . . . . 350

12-6. Sample IV - Within-treatment Differences on California Test of Personality Variables . . . . . . . . . . . . . 352

12-7. Sample IV - Sex Differences on California Test of Personality Variables . . . . . . . . . . . . . . . . 354

12-8. Between-treatment Differences on the GAMIN Ascendancy and Inferiority Scores . . . . . . . . . . . . . . 357

12-9. Within-treatment Differences on the GAMIN Ascendancy Score . . 358

12-10. Within-treatment Differences on the GAMIN Inferiority Score . . 359

12-11. Sex Differences on the GAMIN Ascendancy and Inferiority Scores . 361

12-12. Sample VI - Between-treatment Differences on Class I Variables of the CPI . . . . . . . . . . . . . . . . . 364

12-13. Sample VII - Between-treatment Differences on Class I Variables of the CPI . . . . . . . . . . . . . . . . . 366

12-14. Significant Level Differences on the CPI Variables for Sample VI . 369

12-15. Significant Level Differences on the CPI Variables for Sample VII . 370

12-16. Sample VI - Between-treatment Differences on the Class II Variables of the CPI . . . . . . . . . . . . . . 373

12-17. Sample VII - Between-treatment Differences on Class II Variables of the CPI . . . . . . . . . . . . . . 375

12-18. Between-treatment Differences on the Class III Variables of the CPI 378

12-19. Between-treatment Differences on Class IV Variables of the CPI . 381

12-20. Sex Differences for Sample VI on the California Psychological Inventory . . . . . . . . . . . . . . . . 383

12-21. Sex Differences for Sample VII on the California Psychological Inventory . . . . . . . . . . . . . . . . 385

13-1. TAT and MPT Cards Used in the Utah Study . . . . . . . . 395

13-2. Sample Composition in the Sixth Grade Study . . . . . . . 401

13-3. Between-treatment Differences on the Three Personality Variables Measured in the Fifth Grade Study . . . . . . . . . . 405

13-4. Differences Between the Sexes in the Fifth Grade Study . . . . 406

13-5. Correlations Between STEP and Personality Variables in Fifth Grade Study . . . . . . . . . . . . . . . . . 408

13-6. Between-treatment Differences on the Sixth Grade Study Variables . 410

13-7. Significant Ability Level Differences . . . . . . . . . 413

13-8. Significant Sex Differences on Personality Variables in the Sixth Grade Study . . . . . . . . . . . . . . . . 416

# LIST OF FIGURES

Figure No. Page

1-1. Typical Analysis Pattern Used in the Utah Study . . . . . . . 5

5-1. Comparisons and Results of the California Study Methods Survey . 128

5-2. Comparisons and Results of the "Thinking About Yourself" Test . . 130

5-3. Comparisons and Results of the TAT (Achievement Motives) Test . 132

## Chapter 1

# INTRODUCTION TO THE UTAH STUDY

The purpose of this chapter is to give the reader a brief overall picture of the research project. Topics treated in this chapter are covered in greater detail in subsequent chapters of the report.

## THE PROBLEM

The general goal of this research project has been to study differences in the effects upon elementary, junior high school, and high school pupils of an ability grouping system that differentiates the curriculum principally by adjusting the rate of presentation of curricular materials, and a random grouping system that differentiates the curriculum principally through the use of enrichment. This study explores a wide range of dependent variables that might be affected differentially by the two grouping treatments, including many not previously explored in ability grouping research.

Perhaps the most important variable to be considered in studies involving instructional methods is achievement. This research has explored the question of achievement differences in ability grouped and random grouped classrooms more thoroughly than any work previously published, considering a number of achievement areas over a four-year period, and measuring achievement differences at all grade levels from fourth through twelfth. Data were collected to compare the differences between these grouping treatments at elementary, junior high, and high school level, and to examine separately the achievement differences occurring between superior, average, and slow pupils in the two systems.

Several exploratory studies were also carried out that provide additional information on certain aspects of achievement. One of these studies was concerned with a special superior group at the junior high school level. Another compared underachievers, normal achievers, and overachievers in the two grouping systems.

A careful study was made to determine whether study habits and methods differed under the two grouping treatments. Although claims have been made that ability grouping develops better study methods, no research has been reported to date that explored the differential effects of ability grouping and random grouping upon this variable. This research explored, not only the claim often made by educators that ability grouping leads to better study methods among bright pupils who are not challenged in the random grouping situation, but also analyzed differences and changes in study habits for average and slow pupils in the two systems.

The possible effects of ability grouping upon social adjustment and pupil peer status has also been a cause for concern among educators. This research attempted to answer some of the questions that have been raised through the study of differences in sociometric status between comparable pupils in ability grouped and random grouped classrooms. Among the questions that were explored are: (1) Does ability grouping create a leadership vacuum in groups of average and slow pupils? (2) Does ability grouping reduce the chances of bright pupils to gain social recognition? (3) To what extent do patterns at social recognition differ for superior, average, and slow pupils in ability grouped and random grouped classrooms? Because meaningful sociometric data are much easier to collect in the self-contained classrooms of the elementary schools than in the departmentalized programs found in junior and senior high schools, the data regarding the above questions were collected at the elementary level.

Some educators have condemned ability grouping on the grounds that it is destructive to the individual's self-concept and the development of realistic and desirable aspiration levels. It has been suggested that in the ability grouped classroom bright pupils become arrogant, while slow pupils become insecure and defeated. The slow pupil's low opinion of himself, it has been suggested, leads him to adjust his goals, aspirations, and efforts downward to fit the label placed on him in the ability grouping situation. This research examines, in considerable detail, these and other questions relative to self-concept differences occurring in ability grouped and random grouped classrooms.

Somewhat related to self-concept and aspiration level are a number of personality areas which may also be affected by the ability grouping situation. Personality measures, both of the group administered inventory type and the individually administered projective type, were employed in this research to explore the possible effects of ability grouping upon personality. The areas of aggression, depression, and inferiority feelings were studied in considerable depth while a wide range of other variables that might be affected by the ability grouping situation such as ascendancy, anxiety to achieve, sense of personal worth, and others were explored to some extent.

The attitudes of pupils toward the school environment and their adjustments to it were also explored in this research. This area has received little attention in previous research but appears to be an important consideration in comparing the ability grouping and random grouping treatments. It is generally conceded that the pupil's attitude is an important element in predicting his success in the school situation. As there appears to be several possible ways that ability grouping could affect the pupil's attitude, a measure of pupil attitude was developed and a limited study of pupil attitudes was carried out.

In addition to the dependent variables mentioned above, a number of other questions less directly related to the main goal of the study were explored. Sex differences were examined for a number of the aforementioned variables to determine whether the grouping treatments had a different effect upon the sexes and also to provide additional knowledge about sex differences. Pupils at different ability levels within a given treatment were also compared on many of these variables in order to provide a better understanding of the nature of differences found between the two grouping treatments. Ad hoc hypotheses were developed from some of our early findings that had suggested new areas for exploration. Exploratory studies were designed and carried out to clarify the research results or to follow-up interesting leads emerging from the study.

## RESEARCH PROCEDURES

Two adjacent and closely comparable school districts in Utah provided the setting in which the above questions were explored. One of these districts employed random grouping with enrichment, and the other had adopted a system of ability grouping with acceleration, coincident with the start of this research. Over 2,500 pupils from the fourth, sixth, seventh, eighth, and ninth grade levels were selected in the two districts at the beginning of the study. During the second year of the study this sample was increased to about 4,000 pupils. Research data were collected over a four year period in order to appraise the long term effects of the two grouping treatments. Pupils who were first tested early in grade four were followed through grade seven. As all samples were similarly followed over the period of the study, data were collected at all grades from four through twelve.

The district that employed ability grouping, referred to throughout this project as District A, used a composite achievement test score as a basis for grouping elementary school pupils into self-contained classrooms at the elementary level. At the secondary level, ability grouping was employed in mathematics and science only. Pupils were grouped into three ability levels referred to in this report as superior, average, and slow in both the elementary and secondary schools.

The other district participating in this research, known as District R, employed random or heterogeneous grouping and differentiated the curriculum for individual differences in the random grouped classroom through the use of enrichment. The District A and District R curricula were very similar in most subject areas and at most grade levels, and in many instances the materials employed in District A to adjust the rate of presentation of the curriculum were also used in District R to provide enrichment in heterogeneous classes. This project called for comparisons between pupils in the two grouping treatments who were of comparable ability level. Pupils in District A who had been assigned to the three types of ability groups were measured by the investigators and then, after administering the same measure to District R pupils, the latter were classified into three ability levels so as to provide samples of pupils of comparable ability level in each of the two districts.

Except during the pilot study year, the measures employed in this study were not part of the regular testing program of either district and were administered by teams of examiners trained and supervised by the principal investigator. These examiners were employed by Utah State University and were not affiliated with either school district.

## ANALYSIS

Data collected in this research were analyzed using several procedures. In cases where it was desirable to measure changes that had taken place in pupils who had been exposed to the two grouping treatments, analysis of covariance was used. For example, this technique was employed in studying the achievement data. In this case, the final achievement scores of the various groups studied were adjusted to compensate for differences present at the beginning of the period covered by the analysis. Analysis of covariance was also used to determine whether significant differences were present in such variables as study habits and self-concept, after adjustments had been made for differences present at the beginning of the study.

Analysis of variance was employed in comparing subjects on most of the other variables. Variables that were analyzed using this technique included those in which it appeared reasonable to assume that the sampling techniques would lead to initially comparable groups, and those on which initial measures could not be obtained. An initial measurement of study habits for Sample IV[1] could not be obtained, for example, because no study habits test had been developed for pupils at this grade level. For most of the samples in the Utah study the personality variables could not be measured twice for the same reason. It appears very likely that the assumption of initial comparability that must be made when analysis of variance is used was justified in most cases. Perhaps the best evidence to support this assumption is the fact that when covariance analysis was carried out, it was found in a majority of the analyses that the groups studied were initially comparable on the variables measured and therefore, little or no adjustment of the final means was necessary.

---

[1] Pupils in the fourth grade during the first year of the study.

For other portions of the research data, chi square analysis was employed. This technique was generally used in cases where the data fell into a number of well-defined catagories or in cases where parametric statistics did not appear appropriate. Much of the sociometric data were analyzed using chi square contingency tables because this permitted use of the standard sociometric classifications, (stars, isolates, and so on) that are generally understood by persons familiar with this area of measurement.

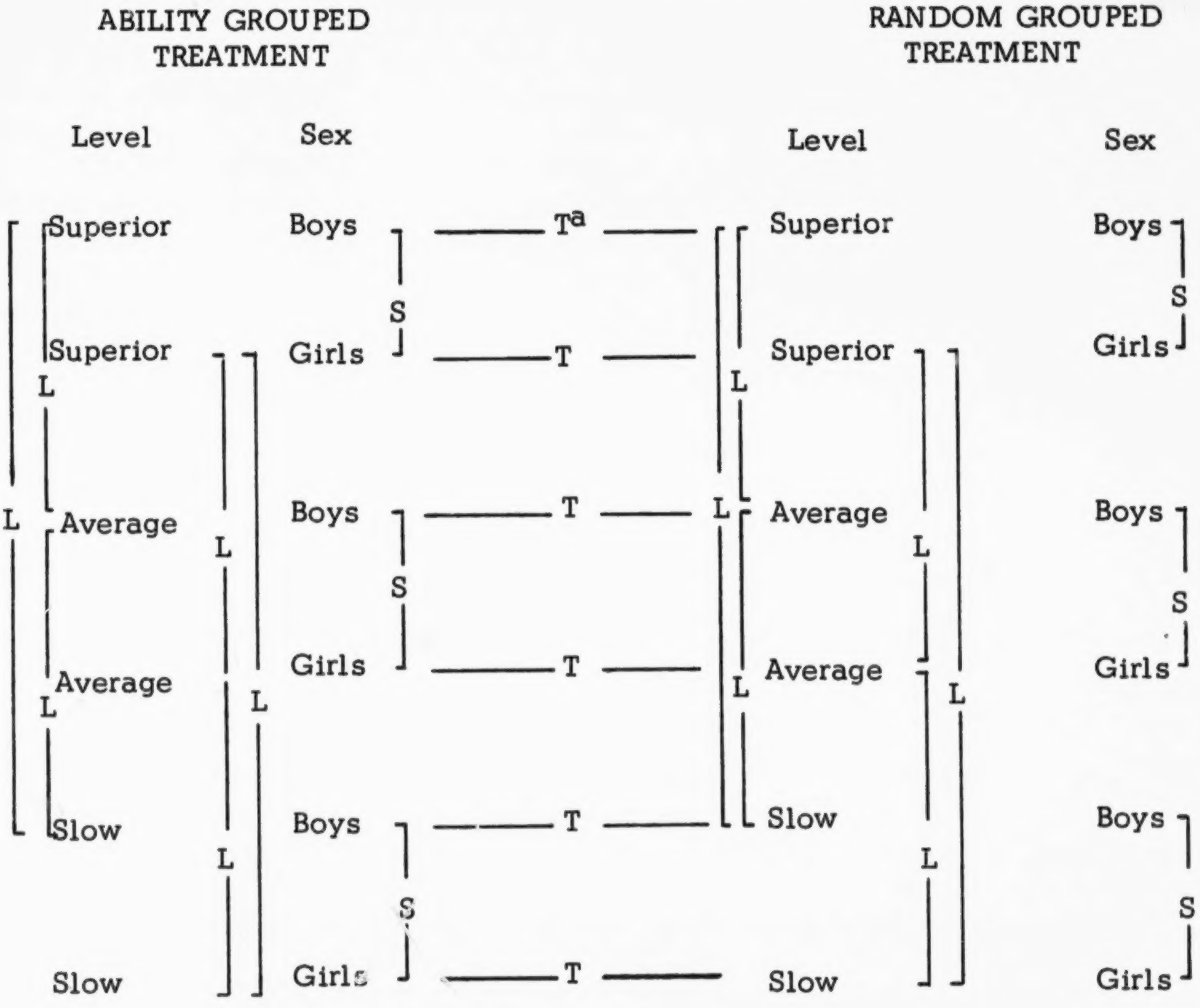
Three major types of differences were considered in analyzing most of the data. The first and most important comparisons were between comparable pupils in the two grouping treatments--ability grouping and random grouping. These comparisons involved pupils of the same sex and ability level in the ability grouped and random grouped samples and are referred to throughout the study as between-treatment differences. These differences provide the best evidence on the differential effects of ability grouping and random grouping upon the dependent variables.

Comparisons between pupils of the same sex and in the same grouping treatment who differed in ability level are referred to as level or within-treatment differences. These comparisons tell us whether the grouping treatments had a differential effect upon pupils at different ability levels and give us additional insights about the treatment that help to interpret the between-treatment differences. The final classification used in the analysis was sex. Boys and girls who were in the same grouping treatment and at the same ability level were compared to provide knowledge about the sex-related differences on the dependent variables. In some instances, these sex differences also provided additional insights into the effects of the grouping treatment.

As a rule, all of the above comparisons were made using single classification analysis of variance or analysis of covariance. Figure 1-1 illustrates the three types of comparisons typically made in the Utah Study. In a typical analysis there were 12 groups (two treatments x three levels x two sexes). Six between-treatment comparisons, twelve level comparisons and six sex comparisons were usually made.

## ORGANIZATION OF THE REPORT

Organizing a research report dealing with a wide range of dependent variables and a number of different samples measured over a four year period presents a number of problems. The usual organization used in reporting less extensive studies was first tried, but was found to be unsatisfactory because of the scope and complexity of this research. The arrangement finally decided upon attempts to organize the study in such a way that all information pertinent to a particular dependent variable would be found in a single chapter. To avoid repeating material pertinent to all of the variables, the report was divided into two major sections. The first section contains information on the sample and the treatments which are applicable to the entire study. The second section contains a separate chapter on each of the major dependent variables with which this research was concerned. Each chapter in the second section presents the


ABILITY GROUPED TREATMENT
RANDOM GROUPED TREATMENT
Level
Sex
Level
Sex
Superior
Boys
Ta
Superior
Boys
S
S
Superior
Girls
T
Superior
Girls
L
L
L
Average
Boys
T
L
Average
Boys
L
L
S
S
Average
Girls
T
Average
Girls
L
L
L
L
Slow
Boys
T
Slow
Boys
L
L
S
S
Slow
Girls
T
Slow
Girls

---

[a] T indicates a between-treatment comparison; L indicates a level or within-treatment comparison; S indicates a sex comparison.

Figure 1-1. Typical analysis pattern used in the Utah Study

working hypotheses relating to the specific variable treated and discusses the rationale underlying these hypotheses, reviews related previous research, describes the tools used to measure the dependent variable, and presents and draws conclusions concerning the results obtained. It is believed that this organization makes the report somewhat more readable and also makes it possible for an individual who is interested in only one or two of the dependent variables covered in the study to obtain a reasonably complete picture by reading only the chapters directly related to his interests.

## Chapter 2

# THE SAMPLE

### INTRODUCTION

As is the case with most field studies in education, random samples of broadly defined populations[1] could not be obtained in the research herein described.

Even if such a random sample had been available, the problem of randomly dividing such a sample and exposing one half of them to a school situation involving ability grouping and the other half to random grouping would have been nearly impossible. Random samples of more narrowly defined populations such as "all pupils enrolled in regular fourth grade classes in school district X" can be obtained in field studies but even this approach often presents administrative problems which public school officials consider too great a disruption of the school's activities. Random samples of narrowly defined populations also have the technical limitation that the results, strictly speaking, are generalizable only to the population from which they were drawn. Thus, the educational field study must usually make compromises in sample selection that would be unacceptable in the more rigorous setting of the psychological laboratory.

Yet, the field study has certain advantages that make it a useful approach in educational research. First, although the rigor of the laboratory is sacrificed, the artificiality that usually attends more rigorous approaches in education is absent or greatly reduced in the field study. Thus, the field study gathers data in realistic and reasonably typical school situations and its findings are often more directly applicable to other school situations. In contrast, the laboratory school which employs highly selected teachers to work with small classes unhampered by financial limitations and free to draw professional support from a university faculty often produces research results that have little carry-over to the usual school in which classes are large, teachers are average, and money is scarce. Second, the field study can work with larger numbers of pupils than are usually manageable in laboratory studies. This permits identifying and studying subgroups that would be too small to consider with the usual laboratory sample. More important, these large samples offer a greater possibility that variables we cannot control, and for that matter, variables that we are not even aware of, will "randomize out" or tend to occur equally in the groups we wish to compare.

The researcher doing a field study, though usually unable to work with truly random samples, is still obligated to make his samples as random as the situation permits and to be alert to biasing elements in his sampling procedure that may distort his findings. His second obligation is to describe with reasonable thoroughness the situation in which the research was carried out. Unless this is done, the reader has little upon which to base his evaluation of the findings of the project and to estimate the pertinence of these findings to other pupils in other schools. The nature of the schools involved, the comparability of their curricula, the numbers of pupils in the sample and how they were selected are among the topics to be discussed in this chapter.

---

[1] Such as "all children enrolled in public elementary schools, grades 4-6, in the United States."

## DISTRICT COMPARISONS

District R from which the random grouped sample was selected, is a city school district with an average daily attendance of about 17,000 (1961-62). District A, from which the ability grouped sample was selected, is the surrounding county school district with an ADA of about 12,500. Similarities and differences between these districts must be considered in interpreting the results of this research. It is impossible, of course, to compare these districts on all aspects of the educational environment they provide and the population they serve. Certain critical comparisons, however, have been made and are reported throughout this chapter.

### Financial Resources

As this study is not concerned with the relative merits of ability grouping and random grouping in an ideal situation but rather in a typical field situation in which financial resources are limited, it was considered advisable to compare the funds available to the two districts to meet their instructional costs. Table 2-1 shows the average amount of money available for instruction per ADA in the two districts. These figures are concerned only with funds directly related to instruction and do not include other operational expenses such as administration, transportation, maintenance, and health services; which, though important, are less directly related to the concerns of this research. It may be seen that District A spent slightly more per pupil during the final 3 years of the study than District R. This difference was greatest in the second year of the study when District A purchased considerable supplemental instructional materials so as to better adapt their curriculum to the three ability levels employed in the grouping program. Program changes, such as establishment of the ability grouping system in District A, usually bring about temporary increases in costs. These increases do not necessarily indicate a proportionate increase in the quality of the program but merely demonstrate that changes cost money. The percentage differences, between the two districts however, are generally small and decreased as the study progressed, which would seem to justify the conclusion that the financial resources available for instruction are reasonably comparable. Thus, we have two districts with similar financial resources to support programs of ability grouping and enrichment and we may appraise from the results of this research the degree to which the two systems achieve certain educational goals and study their relative impact upon the pupil.

### Comparability of Teachers in the Two Districts

Perhaps the most important element to consider in comparing Districts A and R is the relative skill of teachers in the classrooms from which the research samples were selected. Unfortunately, a valid measurement of teaching ability is extremely difficult if not impossible at our present state of knowlege. Thus, our comparisons

Table 2-1

Instructional Budgets Per ADA

| Year | District A | District R | Percent Difference |
|---|---|---|---|
| 1958-59 | $201.58 | $204.14 | 1.27 |
| 1959-60 | $245.12 | $228.84 | 7.11 |
| 1960-61 | $247.49 | $235.74 | 4.98 |
| 1961-62 | $260.84 | $253.12 | 3.05 |

between teachers must rely on variables which, though indirect, are probably valid indicators of the skill of a given group of teachers.

Considerable data were collected during the first year of the research to determine the ability of the two districts to compete with each other for teachers and to estimate the comparability of teachers who participated in the study from the two districts. The salary of participating District A teachers during the first year of the study (1958-59) ranged from $3540 to $5800 with a mean of $4744. This compares closely with District R salaries which ranged from $3250 to $5880 with a mean of $4788.00.

At the elementary level in District A, 84 per cent of the teachers involved in this study had regular Utah Elementary Certificates, 6 per cent had Provisional Elementary Certificates and 10 per cent held Utah Secondary certificates. In District R, 93 per cent held regular Utah Elementary Certificates, 5 per cent held Provisional Elementary Certificates and one teacher was employed under a special state authorization. The overall distribution of experience was similar for the two districts, although the mean amount of teaching experience in District R--17.1 years was greater than in District A--12.6 years. Three teachers in District A were in their first year of teaching during the initial year of the study. In District A, all of the elementary teachers involved in the study held the bachelors degree with 10 per cent also having a Master's degree. In District R, 95 per cent held the Bachelor's degree and 5 per cent had not completed work on the Bachelor's degree.

Among the secondary teachers who participated in the first year of the study, 97 per cent of those in District A and 98 per cent in District R held the regular Utah Secondary Certificate. The remainder held the Utah Provisional Secondary Certificate. The distribution of teaching experience showed that considerably more District R teachers had more than 15 years experience. The mean experience was 6.7 years for District A and 11.5 years for District R. In District A, 17 per cent of the participating secondary teachers were in their first year of teaching as compared with 16 per cent in District R. All of the participating secondary teachers in both districts held the Bachelor's degree with 6 per cent of those in District A and 19 per cent of those in District R also having the Master's degree.

Thus, we see that with respect to most of the above variables, teachers in the two districts were reasonably comparable, with neither district showing a consistent superiority.

## Similarities and Differences in Curriculum

Another important factor that must be considered in comparing Districts A and R is the curriculum. Major differences in the basic curricula of the two districts would make achievement comparisons meaningless as any achievement measure used would then be more appropriate and have higher content validity for one district than the other. A study in which the same basic curriculum were adjusted by the two different treatments--ability grouping with acceleration and random grouping with enrichment--would provide an ideal situation for comparing the treatments. The basic curricula in Districts A and R, though not identical, are very similar and seem to provide a sound basis for comparing the two teaching approaches with which this research is concerned.

Curriculum differences are less critical in comparing other variables such as personality, but at best, would introduce another source of error, the effect of which would be difficult to access.

In the State of Utah, all school subjects at both elementary and secondary levels in which achievement was measured in this research, have curriculum guides that have been developed cooperatively by the Utah State Department of Public Instruction and the school districts of the state. These guides are published by the State Department of Public Instruction and are followed by both of the participating school districts; not only in the subject areas that were measured in this research[1], but also in most other subjects taught in the elementary and secondary schools. Both districts also selected their textbooks from the state approved textbook list which is made up of textbooks meeting the objectives of the state courses of study and are quite comparable. In some subject areas, the same textbooks were employed by the two districts, while in others, different but similar textbooks were employed. Thus, the curricula offered by the two districts were very similar in terms of major objectives and sequences. The state guides, however, permit some latitude and there were specific curriculum differences, particularly at the secondary level, that may have had an effect upon the achievement results in this research. In many cases the basic District A curriculum, which was geared to pupils in the average ability groups, was found to be almost identical in rate, sequence, and content to the District R basic curriculum for heterogeneous groups. Differences did occur, however, between the District A curriculum for superior and slow pupils and the District R curriculum, largely because of the considerable adjustments in rate of progress made in District A. For example, nearly all District A superior pupils were scheduled to take Algebra I in the eighth grade; while District R pupils did not take Algebra I until the ninth grade. As enrichment materials were available in District R classrooms, one would expect that the District R superior pupils, while moving at a slower pace than their counterparts in District A, would have more time for independent study and projects and might develop greater depth in the subjects covered.

## Requirements at Secondary Level

### Junior High School Mathematics

Basic requirements in mathematics were the same for both districts. All pupils were required to take mathematics in the seventh and eighth grades. Ninth grade mathematics was not required in either district but in District A, a course in general mathematics was offered at this level in addition to algebra and pupils were strongly encouraged to take either algebra or general mathematics. As a result, a larger percentage of ninth grade pupils took mathematics in District A than was the case in District R, where a general mathematics course was not offered in the ninth grade.

### Junior High School Science

Both districts required a half-year science course at the seventh grade level. These courses, as offered in the two districts, were very similar in content and objectives although the rate of progress varied considerably with the different ability levels in District A. At the eighth grade level, District A required a full year of science while District R required one-half year. Again, the content and objectives of the courses were similar, placing principal emphasis at this grade level on the life sciences. The longer course required in District A, however, certainly gives this district an advantage at the eighth grade level. This difference should be considered in comparing eighth grade achievement in science. At the ninth grade level, both districts required a one year science course which was essentially an introduction to the physical sciences. This was the course that would be taken by average pupils in District A and by all pupils in District R. Due to the adjustments in rate made in District A, superior pupils would complete this course during the eighth grade, and slow pupils did not reach this point in the sequence.

---

[1] Mathematics, science, reading and social studies at the elementary level and mathematics and science at the secondary level.

## High School Mathematics

At the senior high school level, both districts required one full year of mathematics which could be taken at any of the three grade levels. Pupils taking the minimum high school requirement usually selected courses in either Consumer Mathematics, Algebra I or Plane Geometry depending upon the point reached by the pupil prior to the tenth grade.

## High School Science Requirements

Science requirements for both districts called for a one-year science course to be taken any time during the three years of high school. Both districts offered a wide range of science courses that would meet the requirements such as biology, botany, chemistry, geology, physical science, physics, physiology, and zoology. Because of the wide range of electives and latitude allowed the student in deciding when he would meet this requirement, comparisons of science achievement were difficult to make between the two districts. Data were collected on all pupils to indicate which science courses they had completed and when these courses were taken and comparisons were attempted between pupils of similar ability levels who had taken comparable courses during the same academic year. Due to the large number of courses available and the fact that the requirements could be met any time during the three years of high school, there were not enough comparable students who took the same course during the same academic year to permit a satisfactory comparative analysis in terms of specific course achievement in either mathematics or science. Gross comparisons of the numbers of mathematics and science courses taken by high school samples in the two districts indicated that in terms of overall credits received, the two districts were closely comparable. The lack of sufficient cases for a detailed analysis led to making certain overall comparisons in achievement at the high school level.

## Summary

In summary, it appears that curriculum differences between the districts were small at the elementary school level. At the secondary level, however, some differences in requirements occurred, such as the eighth grade science requirement, which must be considered in interpreting achievement data and certainly make interpretation more difficult. The most serious problem faced in interpreting the secondary achievement data, however, is brought about by the increased number of electives available as we proceed through the grades. This problem first occurred at grade nine and became increasingly important as we advanced through the high school years.

### Selecting Comparable Schools

Many schools in each district, especially at the elementary level, are not comparable to any schools in the other district. Elementary schools in the center of District R are somewhat typical urban schools and have a large proportion of children from the lower-middle and lower classes. In contrast, the more remote county elementary schools in District A are small and predominantly rural. There are, however, a number of elementary schools in the suburban areas that serve very similar populations. Some of these schools are in District A and are in District R. It was primarily from these comparable schools that the initial elementary school sample was drawn. A list of District A and District R elementary schools having at least three sections per grade was drawn up and members of the two central district staffs such as the pupil personnel directors and the elementary supervisors were asked to match schools that they considered comparable in terms of the socioeconomic status and rural-urban-suburban nature of the areas they served. Schools were considered matched when personnel from both districts agreed independently that the schools being considered were comparable on the aforementioned variables.

Selection of the secondary sample that consisted initially of seventh, eighth, and ninth grade pupils was made from junior high schools that served the elementary schools from which the elementary samples had been taken. These schools, however, also served elementary schools not used in the study and therefore were less comparable than the selected elementary schools. To obtain the initial junior high school samples, classes in mathematics and science were selected from these schools at random in District R. In District A, a stratified random sample of classes in these subjects was selected.

## SELECTING RESEARCH SUBJECTS

During the pilot study year of the research project, subjects were selected from the fourth, sixth, seventh, eighth, and ninth grades. Table 2-2 gives the number of pupils in each of these five samples during the four years of the project. These data are based on numbers of pupils completing each achievement test battery. Some phases of the analysis involved larger numbers of subjects than are given in Table 2-2 because some pupils completed other measures administered in the project but missed part of the achievement battery. On the other hand, analyses of measures administered two or more years apart, such as those concerned with achievement gains over a four year period, usually were based upon fewer cases than are shown in Table 2-2. In general, however, this table gives the reader a reasonable picture of the numbers of pupils involved in the Utah study.

Initial samples were drawn in classroom units rather than by individuals. This procedure, of course, is less satisfactory than random selection of individual cases. At the elementary grade levels, it was necessary to include all District A schools with three or more sections per grade in order to obtain enough subjects.[1] In District R elementary schools and in all secondary schools, however, random sampling by individuals would have been possible. The limited funds available during the pilot

---

[1] The fourth grade classes in one available school were not used during the first year, but were added in the second year.

Table 2-2
Number of Pupils in Each Sample[a]

| | Fall 1958 (CAT) | | Spring 1960 (STEP) | | Spring 1961 (STEP)[b] | | Spring 1962 (STEP) | |
|---|---|---|---|---|---|---|---|---|
| | District A | District R | District A | District R | District A | District R | District A | District R |
| Sample IV | | | | | | | | |
| Superior | 100 | 145 | 134 | 253 | 85 | 228 | 82 | 231 |
| Average | 110 | 151 | 228 | 220 | 141 | 159 | 138 | 159 |
| Slow | 83 | 78 | 86 | 110 | 43 | 81 | 43 | 82 |
| Total | 293 | 374 | 448 | 583 | 269 | 468 | 263 | 472 |
| Sample VI | | | | | | | | |
| Superior | 172 | 215 | 168 | 217 | 157 | 195 | 154 | 184 |
| Average | 143 | 140 | 284 | 134 | 258 | 117 | 243 | 115 |
| Slow | 126 | 79 | 126 | 77 | 83 | 82 | 104 | 70 |
| Total | 441 | 434 | 578 | 428 | 498 | 394 | 501 | 369 |
| Sample VII | | | | | | | | |
| Superior | 24 | 72 | 148 | 143 | 110 | 125 | 100 | 120 |
| Average | 47 | 103 | 259 | 129 | 235 | 104 | 214 | 86 |
| Slow | 20 | 38 | 51 | 61 | 55 | 54 | 41 | 42 |
| Total | 91 | 213 | 458 | 333 | 400 | 283 | 355 | 248 |
| Sample VIII | | | | | | | | |
| Superior | 69 | 80 | 110 | 151 | 104 | 121 | 97 | 115 |
| Average | 78 | 57 | 216 | 71 | 185 | 54 | 161 | 55 |
| Slow | 32 | 61 | 54 | 72 | 39 | 51 | 37 | 42 |
| Total | 179 | 198 | 380 | 294 | 328 | 226 | 295 | 212 |
| Sample IX | | | | | | | | |
| Superior | 33 | 92 | 33 | 92 | 27 | 75 | 23 | 65 |
| Average | 88 | 38 | 85 | 38 | 73 | 30 | 69 | 26 |
| Slow | 15 | 41 | 14 | 41 | 13 | 30 | 12 | 24 |
| Total | 136 | 171 | 132 | 171 | 113 | 135 | 104 | 115 |
| Yearly Total | 1140 | 1390 | 1996 | 1809 | 1608 | 1506 | 1518 | 1416 |

[a] Based on achievement test batteries as indicated; California Achievement Tests (CAT) or Sequential Tests of Educational Progress (STEP)

[b] No achievement battery was administered at secondary level this year; Sample VI, VII, VIII and IX based on the Index of Adjustment and Values.

study year for test administration and the reluctance of the schools to cooperate to the greater degree necessary to identify and test a random sampling of individuals led the principal investigator to use of the randomly selected classroom unit as a basis for initial selection of all research subjects during the 1958-59 school year.

Sample IV

Pupils who were in the fourth grade at the start of the project, (1958-59) are referred toas Sample IV throughout this report. This fourth grade sample in District A came from nine classrooms in three elementary schools. Three superior sections, three average sections and three slow sections were included, one of each from each school. Two of these schools had more than one average section. In these cases only one average section was chosen at random from the two available at the school. As the district policy called for pupils to be randomly assigned to average sections if more than one was required, it did not appear necessary to test all average sections as this would have given a somewhat larger sample than the available samples of superior and slow pupils.

The fourth grade sample in District R came from 13 classrooms in six schools. These schools were selected for comparability with the District A schools in terms of socioeconomic status of the neighborhood served. Classrooms were selected at random from among the available classrooms in each school. As District R pupils were randomly grouped, a larger sample was considered necessary in order to insure getting enough subjects at each ability level. At the time classrooms were selected, data were not available to determine the number of pupils in the random grouped classrooms who would be equivalent to pupils at each of the three ability levels in District A. Much of the planned analysis called for breaking down the sample into ability levels, and since considerable loss of subjects may be expected in a four year study, it was considered important to insure sufficient cases at each ability level. During the first year of the study, a total of 394 fourth grade pupils in District R and 303 in District A completed the first achievement battery.

Prior to the beginning of the second year of the study, support was obtained from the Cooperative Research Program, U.S. Office of Education, to carry out the remaining three years of the project. At this time it was decided to increase the number of subjects in the fourth grade sample. A total of 583 pupils in District R and 448 pupils in District A were tested during this year.

At the start of the third year of the project, this sample suffered a serious loss when one of the four elementary schools in District A changed to a two-level ability grouping system because of a drop in enrollment. The principal investigator concluded that grouping at two levels changed the situation sufficiently so that the only course open was to eliminate these subjects from the sample. This loss caused the District A sample during the third and fourth years to be considerably smaller than was anticipated. Very thorough makeup procedures were used during the third and fourth years but a few cases were lost because of prolonged absense during the period when the testing and makeup program was underway. It will be noted in Table 2-2, however, that the number of pupils who actually completed the achievement battery each year in District R increased from year three to year four of the project. As no new subjects were added to the sample during these years, this reflects the effort made to obtain data on as many of the sample as possible.

Some phases of the research employed special sub-samples and in some instances a pupil was included in one part of the analysis who did not have all scores needed for another part of the analysis. Thus, the number of subjects fluctuates in different phases of the study and for different variables with the achievement analysis generally being based on fewer cases than analysis of the other dependent variables such as study habits, self concept, and personality that were based on shorter tests and were easier to make up if missed. Some fluctuations in both groups were due to pupils changing ability level from year to year. Table 2-3 shows the distribution of grade placement scores at the three ability levels for the District A fourth grade sample during the pilot study year. Although the mean achievement levels are markedly different for each of the three ability groups, it may be seen that considerable overlapping is present.

### Sample VI

Pupils sampled from the sixth grade during the initial year of the study have been designated as Sample VI. The District A Sample VI was taken from 12 classrooms in four schools and included four superior sections, four average sections and four slow sections. The procedure for selecting sixth grade classrooms was the same as for fourth grade classrooms in District A. Table 2-4 shows the distribution of the three District A ability groups on the first achievement test battery. More overlapping occurs between combined ability groups than between ability groups in any one school as criteria for grouping were not absolute but instead were based on the relative achievement within each school. The sixth grade sample in District R was taken from 16 classrooms in eight schools following the same procedure employed in selecting fourth grade classrooms in District R.[1] Table 2-2 summarizes the total number of pupils who completed the achievement test batteries during the 1958-59, 1959-60, and 1961-62 school year. As no achievement test battery was administered to this sample or other secondary samples in the 1960-61 school year, the figures for that year represent the number of pupils completing the Index of Adjustment and Values.[2]

### Sample VII

Pupils enrolled in the seventh grade during the initial year of the study who were selected as part of the research sample are designated as Sample VII. It may be seen that during the pilot study year, the District A sample consisted of all pupils in three seventh grade mathematics classes, one each being randomly selected from the superior, average and slow classes. This sample was considerably enlarged during the following year as may be seen in Table 2-2. Subsequent to the 1959-60 year, no new cases were added to Sample VII. The Sample VII subjects for District R during the first year were taken from seven randomly selected seventh grade mathematics classes. This sample was also substantially enlarged during the second year of the study.

---

[1] A total of 18 classrooms were tested but two were randomly eliminated when it was found that sufficient cases were available at each ability level using only 16 classes.

[2] This measure was used because it was administered all secondary school samples during the 1960-61 year.

Table 2-3

Distribution of District A Elementary Sample,
4th Grade on CAT First Battery.
Total GP Score

| | Slow<br>4th grades<br>75, 81, 87 | | Average<br>4th grades<br>76, 83, 88 | | Superior<br>4th grades<br>77, 82, 89 | |
|---|---|---|---|---|---|---|
| Composite GP Score | Boys | Girls | Boys | Girls | Boys | Girls |
| 60 - 64 | | | | 1 | 2 | 3 |
| 55 - 59 | | | 1 | 3 | 9 | 16 |
| 50 - 54 | | 1 | 2 | 2 | 13 | 20 |
| 45 - 49 | 1 | | 14 | 12 | 17 | 14 |
| 40 - 44 | 13 | 4 | 24 | 17 | 1 | 11 |
| 35 - 39 | 18 | 11 | 13 | 11 | 2 | |
| 30 - 34 | 13 | 10 | 8 | 2 | | 2 |
| 25 - 29 | 8 | 1 | | | | |
| 20 - 24 | 3 | | | | | |
| Totals | 56 | 27 | 62 | 48 | 34 | 66 |
| Means | 3.53 | | 4.23 | | 5.06 | |

Table 2-4

Distribution of District A Elementary Sample,
6th Grade on CAT First Battery.

| | Slow<br>6th grade<br>78, 84, 90, 93 | | Average<br>6th grade<br>80, 85, 91, 94 | | Superior<br>6th grade<br>79, 86, 92, 95 | |
|---|---|---|---|---|---|---|
| Composite GP Score | Boys | Girls | Boys | Girls | Boys | Girls |
| 80 - 84 | | | | | 1 | |
| 75 - 79 | | | | | 6 | 5 |
| 70 - 74 | | | | 1 | 16 | 25 |
| 65 - 69 | | 2 | 10 | 5 | 25 | 41 |
| 60 - 64 | 4 | 3 | 22 | 31 | 12 | 11 |
| 55 - 59 | 12 | 11 | 22 | 23 | 6 | 1 |
| 50 - 54 | 16 | 19 | 11 | 6 | 2 | |
| 45 - 49 | 25 | 10 | 3 | 3 | 1 | |
| 40 - 44 | 9 | 4 | | 1 | | |
| 35 - 39 | 11 | 1 | 1 | 1 | | |
| 30 - 34 | 1 | | | 1 | | |
| 25 - 29 | | | 2 | | | |
| Totals | 77 | 49 | 71 | 72 | 69 | 103 |
| Means | 4.94 | | 5.80 | | 6.76 | |

## Sample VIII

Eighth grade pupils selected during the pilot study year are designated as Sample VIII. These pupils, as was the case with Sample VII, were taken from the junior high schools which ordinarily serve the schools from which the elementary sample was drawn. District A Sample VIII consisted of all pupils in six randomly selected eighth grade mathematics classes. In District R, Sample VIII consisted of all pupils in seven randomly selected eighth grade mathematics classes. Both District A and District R samples were enlarged during the second year of the study (see Table 2-2).

## Sample IX

Pupils in the ninth grade during the initial year of the study were designated as Sample IX throughout the course of the study. These pupils were studied from the beginning of the ninth grade through the end of the twelfth grade. In District A, this sample included all pupils in three ninth grade algebra classes, one each being randomly selected from low ability, average ability and superior ability classes available. The remainder of Sample IX from District A was taken from three ninth grade science classes, one each at slow, average, and superior levels. As many pupils at the superior and slow levels were in both the mathematics and science classes selected, this sample is quite small.

In District R Sample IX included all pupils taken from four randomly selected ninth grade science classes and four randomly selected ninth grade mathematics classes. This sample tended to yield a larger number of superior and average pupils and a smaller number of slow pupils because of the fact that ninth grade algebra is not as strongly recommended to pupils in District R and is taken by a relatively small number of slow pupils. The main effect of this error upon the results was to reduce considerably the number of slow pupils available for analysis. As comparisons between treatments were made within ability levels, i.e., District A superior pupils were compared to District R superior pupils, and overall comparisons were not generally made, a change in the proportion of slow and superior pupils in the sample does not distort the research findings. Table 2-2 shows the numbers of Sample IX pupils who completed the achievement battery in 1958-59, 1959-60, and 1961-62. As was the case for other secondary school samples, 1960-61 figures are based upon the number of Sample IX pupils completing the Index of Adjustment and Values.

## Sample Variations

Because a fairly large number of variables were explored in this research, the numbers of pupils for whom complete data are available changes from one variable to another. For example, a pupil may have completed all tests relating to personality or self-concept but may have missed and been unable to make up one of the achievement tests. He may, therefore, be included in the self-concept and personality analysis but not in parts of the achievement analysis. The number of students available for achievement analysis is generally smallest because the achievement batteries were longer, were usually administered over a period of four days at the elementary level and two days at the secondary level and were thus more difficult to make up.

## COMPARABILITY OF RESEARCH SUBJECTS

Although considerable effort was made in selecting the samples to randomly select from the population and avoid techniques that would lead to biasing of the samples, the techniques used were far from ideal. Thus, during the course of the study, and especially in the pilot study year, data were collected to check the success of the techniques used in selecting reasonably comparable subjects. Some of these data will now be presented.

### Socioeconomic Class Differences Between District A and District R Pupils of Comparable Ability

It will be recalled that elementary schools were selected at the outset of the study that were considered to serve comparable socioeconomic groups in the two districts. A preliminary check of socioeconomic status made in these schools indicated no significant differences. It was considered desirable, however, to make a more thorough check of socioeconomic differences as this variable could have a considerable effect upon the outcome of the research. The Hollingshead Occupational Scale was selected as a basis for establishing five socioeconomic class levels. This scale is based entirely upon the occupation of the father or principal family wage earner. Although much easier to obtain than some of the more elaborate socioeconomic class measures such as Warner's Index of Status Characteristics, the Hollingshead Occupational Scale has been found to correlate highly with such measures (1958). The investigator's previous experience has shown that considerable information is required on a written questionnaire in order to obtain sufficient data to make a valid occupational classification. Therefore, a biographical information form was developed for this study which included descriptions and examples of the five job levels used in Hollingshead's classifications. The pupil was asked to select the classification best fitting the work his father (or the principal family wage earner) did. He was also asked to tell where his father worked, give the title that his father had in his work and to write a description of the work done. It was possible, by cross checking the responses to these different questions, to classify nearly all of the occupations without seeking further information. In a few cases, discrepancies in responses to the different questions or inadequate answers led the investigator to check further with the school counselor, teacher, or the parents. These data were collected during the final year of the study. At this time, Sample IV and Sample VI pupils were in the seventh and ninth grade respectively. At this level they possessed more information on their father's occupations than could have been obtained at the elementary school level.

The results of the analysis of the occupational data are shown in Table 2-5. The distribution of cases into the five Hollingshead socioeconomic class levels was compared using chi square contingency tables to determine whether pupils of comparable ability levels in the two districts were significantly different in their distribution into the five social class levels. Of the six chi square values calculated between pupils of comparable ability in Samples IV and VI, one was statistically significant. The significant chi square was found between District A and District R superior pupils in Sample VI. This chi square reached significance at the .01 level and indicated a trend towards District R superior pupils being of higher socioeconomic status. None of the other chi square values reached significance at the 5 per cent level and most

Table 2-5

Socioeconomic Class Differences Between District A and District R Pupils of Comparable Ability

| Group | CAT. 1 Percent | CAT. 2 Percent | CAT. 3 Percent | CAT. 4 Percent | CAT. 5 Percent |
|---|---|---|---|---|---|
| | Percent of Pupils in the 5 Hollingshead Scale Categories | | | | |
| District A -- Superior -- IV | 5 | 16 | 25 | 44 | 10 |
| District R -- Superior -- IV | 10 | 25 | 26 | 28 | 11 |
| District A -- Average -- IV | 3 | 8 | 18 | 52 | 19 |
| District R -- Average -- IV | 9 | 15 | 23 | 36 | 17 |
| District A -- Slow -- IV | 0 | 7 | 20 | 48 | 25 |
| District R -- Slow -- IV | 4 | 10 | 19 | 48 | 19 |
| District A -- Superior -- VI | 3 | 22 | 21 | 42 | 12 |
| District R -- Superior -- VI | 10 | 31 | 24 | 21 | 13 |
| District A -- Average -- VI | 4 | 11 | 15 | 48 | 22 |
| District R -- Average -- VI | 3 | 18 | 28 | 37 | 14 |
| District A -- Slow -- VI | 2 | 2 | 21 | 52 | 23 |
| District R -- Slow -- VI | 5 | 9 | 18 | 48 | 21 |

were small, indicating very slight differences in socioeconomic distribution between the two samples. The significant difference, however, between superior Sample VI pupils in Districts A and R should be taken into consideration in interpreting differences found between these groups on the other variables considered in this research. It should be noted, however, that data obtained from a random subgroup of Sample VI pupils during the second year of the study showed no significant differences between the two districts. These data were obtained by individual interview and were based on a weighted combination of the *Hollingshead Occupational and Educational Scales* (see Chapter 5).

## Intelligence Differences at Beginning of the Study

Mean intelligence levels for groups of comparable ability in District A and District R are given in Table 2-6. It may be seen that most of the parallel groups in the two districts were closely comparable in mean mental age as expressed by intelligence grade placement equivalents. The District A data in this table also demonstrates that use of the composite score on the *California Achievement Test* as a basis for grouping differentiated the District A groups successfully in terms of intelligence or scholastic aptitude as measured by the *California Short-Form Test of Mental Maturity* (CMM). It will be noted that the language score differentiates the three ability groups somewhat better than the non-language score. For the language score on the *California Short-Form Test of Mental Maturity* the average difference between comparable ability groups in Districts A and R is less than two-tenths of a grade. The greatest difference was in Sample IV and favored the District R slow sample.

The differences between comparable District A and District R pupils were somewhat larger on the *CMM* non-language scores, averaging nearly three-tenths of a grade. The greatest difference was found between the slow pupils in Sample IX and favored the District R pupils. Even this largest difference, represents less than two raw score points (28.95 vs. 27.25) and falls far short of statistical significance with a t-ratio of 1.29. For comparable groups in the two districts where the difference in mean intelligence grade placement exceeds three-tenths of a grade, this variable will be considered in interpreting achievement results. It should be noted, however, that language I.Q. is usually found to be more closely related to scholastic success than non-language I.Q., and is thus the more pertinent variable in an ability grouping study.

## REFERENCES

Hollingshead, A. B., Redlich, F. C. *Social class and mental illness*. New York: Wiley, 1958.

Table 2-6

CMM Language and Non-Language
Intelligence Grade Placements[a]

| | District A | | District R | |
|---|---|---|---|---|
| Group | Lang. | Non-Lang. | Lang. | Non-Lang. |
| Sample IV (4th grade) | | | | |
| superior | 5.7 | 5.4 | 5.6 | 5.2 |
| average | 4.5 | 4.7 | 4.5 | 4.3 |
| slow | 2.9 | 3.9 | 3.6 | 3.7 |
| Sample VI (6th grade) | | | | |
| superior | 8.3 | 7.1 | 8.1 | 7.0 |
| average | 6.7 | 6.1 | 6.5 | 6.1 |
| slow | 5.2 | 5.4 | 5.3 | 5.2 |
| Sample VII (7th grade) | | | | |
| superior | 9.0 | 8.3 | 9.0 | 8.4 |
| average | 7.1 | 8.0 | 7.3 | 7.6 |
| slow | 5.6 | 7.0 | 5.9 | 6.5 |
| Sample VII (8th grade) | | | | |
| superior | 9.8 | 9.0 | 9.7 | 8.5 |
| average | 8.0 | 8.0 | 7.7 | 7.7 |
| slow | 6.1 | 7.0 | 6.0 | 6.6 |
| Sample IX (9th grade) | | | | |
| superior | 10.4 | 9.0 | 10.4 | 9.3 |
| average | 9.0 | 8.0 | 9.1 | 8.2 |
| slow | 7.7 | 7.1 | 7.3 | 7.7 |

---

[a] Conversion tables used will be found in: Sullivan, E.T., W.W. Clark, and E.W. Tiegs. Manual California Short-Form Test of Mental Maturity, 1957 S.-Form (Elementary and Junior High Levels) Los Angeles: California Test Bureau, 1957.

# Chapter 3

# THE TREATMENTS

## INTRODUCTION

Because terms such as "ability grouping," "enrichment" and "heterogeneous grouping" have different meanings in different situations, it is necessary to define operationally their meaning in a research project so that the reader fully understands the specific circumstances under which the data were collected. Many of the conflicting results found in different educational research projects may be traced to studies which appear on the surface to be concerned with the same treatments but in reality are not. In ability grouping studies, for example, the term sometimes means grouping on the basis of a general criterion such as an intelligence test, sometimes grouping on the basis of teacher judgment, and sometimes grouping on the basis of specific achievement in a particular subject, such as mathematics. It sometimes means grouping within a single classroom, sometimes grouping all pupils at a given grade level and sometimes grouping in ungraded classes which may cut across several grade levels. Let us now examine the systems for grouping pupils employed by District A and District R in this research so as to provide the reader with operational definitions of the specific nature of the two grouping treatments that constitute the independent variable in this research.

## THE ABILITY GROUPING SYSTEM

The ability grouping system used in District A during the course of this research employed a standard achievement battery, the California Achievement Test Battery (CAT), as a basis for grouping pupils into three ability levels. An achievement measure was selected instead of scholastic aptitude or intelligence measures largely because the district administration found that grouping on the basis of past achievement was more acceptable to parents than grouping on the basis of ability. Scores on a battery of achievement tests have generally been found to correlate rather highly with measures of scholastic aptitude, so it is doubtful if any marked difference in the grouping would grow out of a change from one of these measures to another.

### Elementary

At the elementary school level, the total grade placement score on the California Achievement Test Battery, form w, was used as a basis for grouping. This battery includes tests in reading, arithmetic, and language. The procedure used is described in detail in Appendix A. Pupils were administered the test battery in the spring of each school year. Pupils at a given grade level were then divided into boys and girls and a separate list made up for each sex starting with the pupil having the highest grade placement score and proceeding in order to the pupil with the lowest score.

The lowest ten boys and lowest ten girls were then selected and were placed in the slow section. Next, the highest 15 boys and highest 15 girls were selected to make up the superior section. The remaining pupils were placed in two or three average sections by placing every second or third pupil in a given section. Thus, the average sections within a given school were comparable in terms of ability level and range rather than being divided into "high average" and "low average" sections. Both districts had a separate program of special classes for mentally retarded pupils, so these are not included in either the ability grouped or random grouped classes.[1] It may be seen in Appendix A that a number of special provisions are made for unusual situations such as an excessive number of superior or slow pupils in a given school, too large an ability gap between boys and girls assigned to a slow section, and others that seemed to call for special treatment.

Thus, at the elementary school level a simple three level system based on a composite achievement score was employed in District A. This system has both advantages. Among its advantages are: first, it may be carried out in relatively small schools, i.e., those with three or more classrooms in the grades to be grouped. Second, the grouping criterion is easy to use and is generally accepted by parents. Third, by grouping within grade levels only and by grouping on a general criterion, this system permits a relatively easy transition for pupils transferring to or from schools employing random grouping. Finally, this system permits the use of self-contained classrooms. When specific grouping criteria are established on a subject by subject basis, some sort of departmentalized program is usually called for; as pupils will sometimes fall in different ability groups for different subject areas.

The principal disadvantage of the grouping system employed in District A is that the use of a general grouping criterion such as a composite achievement test score results in only a moderate reduction in heterogeneity. It does not produce pupils who are at or even near the same level on all subjects included in the curriculum. Thus, at the elementary school level, District A employed what might be referred to as a moderate grouping system which can be adapted to most elementary schools, but which brings about only moderate reductions in heterogeneity.

### Secondary

At the secondary school level, District A employed ability grouping in mathematics and science classes only. Even within these two subject areas, the classes that could be satisfactorily grouped became fewer as the pupil moved into the higher grade levels. At these levels, a large number of elective courses and relatively few specific requirements result in few courses in which a sufficient number of sections are available to permit grouping on ability. The grouping criterion for mathematics courses was the grade placement score of the mathematics subtest of the California Achievement Test Battery. The achievement score obtained in the spring on this subtest formed the basis for grouping pupils in mathematics classes the following fall. In science, on the other hand, the composite grade placement score for the Mathematics and Language Subtests of the California Achievement Test battery was used, thus providing a general criterion for grouping in this subject. The use of a general criterion in science and a specific criterion in mathematics permits us to compare the relative outcomes of the two types of criteria on a number of variables. In District A, ability grouping was employed in all mathematics and science classes during the junior high school grades.

---

[1] Some evidence gathered in the study suggests that mentally retarded pupils may function in a satisfactory manner in a slow, ability grouped class (see Chapter 13) and it is possible that some mentally retarded children were included in the District A elementary school samples.

At the high school level, on the other hand, ability grouping was only possible in the lower level mathematics and science courses taken by a sizeable percentage of pupils in the schools. In such elective courses as physics and chemistry, the maximum amount of grouping possible was a two level grouping system as virtually no slow pupils took these elective physical science courses.

## THE ENRICHMENT SYSTEM

In District R, a program of heterogeneous grouping plus enrichment was employed at all grade levels. At the elementary school level, pupils were randomly assigned to the various sections within a given grade level. All principals in District R were briefed on the research project by the project director during the pilot study year of the project and the Superintendent in District R set up a specific policy for placing pupils in sections which required that pupils be randomly assigned. At the secondary level, the same policy was in effect. Certain factors, however, led to reducing the variability of District R pupils in terms of ability in certain classes at the secondary level, with this trend becoming more pronounced in the advanced grades. One factor that leads to this grouping is the limited mathematics and science requirements in both districts--one year of each to be taken any time during the three years of high school. Another factor is the wide variety of electives available to the student. In District R, as is the case in most schools, a considerably larger percentage of superior pupils go on to take mathematics and science courses in excess of the minimum requirements while relatively few slow pupils proceed beyond the minimum requirement.

Another factor operating to bring about a reduction in class heterogeneity in District R is the relationship between ability and certain electives such as foreign language and band. If an elective such as band is offered only during period five, the science classes also offered during period five will probably have fewer than the usual proportion of superior pupils because many of the superior pupils will be enrolled in band and will, therefore, have to take their mathematics and science courses during some other period.

## REDUCTION IN HETEROGENEITY IN DISTRICT A

### Introduction

Because ability grouping systems attempt to achieve a more favorable situation for curriculum adjustment by reduction of heterogeneity within the classroom, it is important in the appraisal of such a system to determine just how much of a reduction in heterogeneity is achieved by the system. The amount that heterogeneity can be reduced by ability grouping is dependent upon several variables. Among these are the spedificity and validity of the grouping criterion, the number of grouping levels to be established and the variability and numbers of the pupils to be grouped. When we have large numbers of heterogeneous pupils, employ multilevel grouping, and have a highly valid criterion, we can achieve great reduction in heterogeneity in a specific subject.

At one extreme, a grouping system could be developed that would group a large number of pupils at a single grade level into a dozen or more different ability levels in a specific subject such as eighth grade arithmetic for which valid grouping criteria are available. Such a system would produce highly homogeneous groups and require few curriculum adjustments within a given group. At the other extreme, we might take the one room school as an example. Here we have a few pupils spread over perhaps eight grades and varying greatly in ability and past achievement. In such a situation, we have a highly heterogeneous group and, as interclass grouping is not possible, great adjustments for individual differences within the group must be made.

The usual ability grouping system attempts to achieve moderate reductions in the heterogeneity of the classroom group while using procedures that can be applied to all but the very small schools. Such systems usually employ three to five grouping levels and use a general criterion such as I.Q. at the elementary level and specific criteria such as achievement or aptitude tests in specific subjects at the secondary level.

A criticism often leveled at ability grouping programs, particularly those employing general criteria such as an intelligence test score or a composite achievement test score has been that such a program reduces the heterogeneity of pupils in specific subjects to such a slight degree that the teacher in effect has just as heterogeneous a class in specific subject areas as would be found if no grouping had taken place. We will now review data concerned with the heterogeneity of District A and District R classes to make some appraisal of the reduction of heterogeneity taking place in District A as a result of the ability grouping system.

## Reduction Achieved

During the pilot study year, the samples at the various grade levels were studied carefully to determine the reduction in heterogeneity achieved by the grouping system used in District A. Distributions of scores found in Tables 2-3 and 3-1 for the fourth grade sample and Tables 2-4 and 3-2 for the sixth grade sample show that some reduction in heterogeneity was achieved although considerable overlapping was present. For District A, classes of the same ability level classification from different schools are combined in Tables 2-3 and 2-4 which tends to increase the apparent overlapping. Even if classes of different ability levels were set up in a single school without overlapping, when different schools were combined overlapping would appear because of different ability cutoffs at different schools. The overlapping during the first year, however, was partially due to different interpretations of the district's grouping directive, which led District A personnel to establish more specific rules for grouping (see Appendix A) which would be expected to have reduced overlapping during the final three years of the study.

In order to determine the variability of classrooms included in the research in District A and District R, means and standard deviations were computed for each elementary school classroom on the California Mental Maturity Short Form Language Score, the CAT Arithmetic total, CAT Language total, and the CAT Reading total. These data may be found in Appendix B. Data on the variability of the three subtests of the California Achievement Battery are summarized in Table 3-3. The average variability in District A on the various subtests was found to be 26 per cent less than in District R for the elementary sample. The smallest difference in variability was found between the arithmetic subtest at the fourth grade level where the difference in average sigma is less than 10 per cent and the greatest variability is found in the sixth grade reading score where the difference in sigma is nearly 45 per cent between District A and District R.[1] As might be expected in a study involving as many different classrooms as are concerned in this study, some of the most homogeneous of the random grouped classrooms have lower standard deviations on some of the measures than the more heterogeneous of the ability grouped classrooms. In cases where low variability was found among the District R classrooms, a check was made to determine whether any deliberate grouping had been made. In each case it was established that the low variability had occurred by chance.

A similar reduction in variability was found at the junior high school level. At this level, however, the reduction in variability was usually much greater for superior and slow groups than for average groups. This is because a larger proportion of the pupils were placed in average groups at the junior high school level and therefore, these average groups would be expected to be nearly as variable as random groups in District R. It may be seen in Appendix B that at each grade level the most heterogeneous groups in District A are more variable than the most homogeneous groups in District R. This illustrates that considerable chance fluctuation in degree of homogeneity takes place in random grouped classrooms. A more rigid system of ability grouping employing more than three levels would increase the homogeneity of classroom groups but would be less broadly applicable, especially at the elementary school level. Studies such as those reported by Worlton (1928), French (1960), and Goldberg et. al. (1961) have demonstrated that class homogeneity without specific curriculum adjustments for class level has little effect on achievement. District R pupils in classes that were more homogeneous may have performed somewhat differently than did pupils in typical random grouped classes. As the District R goal is enrichment while the District A goal is acceleration, it is doubtful that the differences in heterogeneity found among District R classes made a significant change in teaching approach in these classes. If sufficient cooperation could be obtained from the schools, perhaps a better design would involve building heterogeneity into control group classes through stratified sampling, just as homogeneity is built into ability grouped classes.

As reduction of heterogeneity in District A could only be determined in the three areas of the CAT battery, the fact that the CAT composite score from the previous year served as the main criterion for grouping, would certainly lead to a greater apparent than real reduction in heterogeneity. It is safe to assume that the ability grouping system in District A resulted in much less reduction in heterogeneity in areas not measured on the CAT battery such as art, music, and social studies.

## CURRICULUM ADJUSTMENTS

### Introduction

In the preceeding section we have discussed the procedures used by Districts A and R to group pupils for instruction and the heterogeneity of classes under these systems. Grouping systems, however, are but a means to an end--this end being to provide a situation in which better learning can take place.

---

[1] Such averages have limited meaning but may be taken as a rough approximation of the degree of reduction in heterogeneity in specific areas.

### Table 3-1

### Distribution of District R Elementary Sample, 4th Grade on CAT First Battery

| GP Score[a] | Teacher Code Number | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 23 | | 28 | | 29 | | 30 | | 34 | | 35 | | 38 | |
| Sex | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| 65 - + | | | | | | | | | | | | | | |
| 60 - 64 | | | 1 | | | | 1 | | | | | 2 | | 1 |
| 55 - 59 | 1 | | | 1 | 2 | | 2 | 4 | 1 | 1 | 1 | 1 | 3 | 2 |
| 50 - 54 | 1 | 3 | 3 | 4 | 2 | 4 | 1 | 2 | 3 | 1 | 2 | 3 | | 3 |
| 45 - 49 | | 4 | 8 | 5 | 7 | 4 | 3 | 6 | 1 | 3 | 4 | 4 | 1 | 2 |
| 40 - 44 | 2 | 7 | 4 | 1 | 4 | 3 | 5 | 3 | 3 | 6 | 2 | 7 | 1 | 2 |
| 35 - 39 | 2 | 4 | 2 | 2 | 3 | | 1 | | 2 | 2 | 2 | | 2 | |
| 30 - 34 | 2 | | 2 | | 1 | 2 | | | 2 | 1 | | 1 | | |
| 25 - 29 | | | | | 1 | | 1 | | | | | | | |
| Totals | 8 | 18 | 20 | 13 | 20 | 13 | 14 | 15 | 12 | 14 | 11 | 18 | 7 | 10 |

[a] Grade-placement score

| 41 | | 42 | | 43 | | 48 | | 49 | | 50 | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | |
| | | | | | | 1 | | | | | 1 | 2 |
| | 1 | | | | | | 1 | | 2 | | 1 | 10 |
| 2 | 2 | 2 | 1 | | | 5 | 2 | 1 | 1 | 2 | 1 | 38 |
| | 7 | 4 | 3 | | 1 | 4 | 7 | 3 | 4 | 2 | 3 | 70 |
| 9 | 2 | 4 | 4 | 4 | 7 | 5 | 7 | 5 | 1 | 5 | 3 | 108 |
| 4 | 5 | 5 | 3 | 5 | 3 | 3 | 3 | 6 | 2 | 3 | 1 | 93 |
| 2 | 1 | 2 | 4 | 6 | 2 | 4 | 2 | | | 1 | 2 | 48 |
| 1 | | 1 | 1 | 1 | 1 | 2 | 1 | 1 | | 1 | | 21 |
| 18 | 18 | 18 | 16 | 16 | 14 | 24 | 23 | 16 | 10 | 14 | 12 | 392 |

Table 3-2

Distribution of District R Elementary Sample, 6th Grade on CAT First Battery Total GP Score

| GP Score | Teacher Code Number | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 26 | | 27 | | 31 | | 32 | | 33 | | 36 | | 37 | | 39 | | 40 | |
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| 85-89 | | | | | | | | | | | | | | | | | | |
| 80-84 | | | | | 1 | | | | | | | | | | | | | |
| 75-79 | | | | | | | | 2 | 1 | 1 | | 1 | | | 1 | 2 | | |
| 70-74 | 3 | 1 | | | 8 | 3 | 1 | 1 | 4 | 1 | | 1 | | 1 | 1 | 1 | | 1 |
| 65-69 | 1 | 5 | | | | 3 | 2 | 1 | 4 | 4 | | 3 | 1 | 4 | 2 | 3 | 4 | 3 |
| 60-64 | 4 | 4 | | | 3 | 4 | 5 | 7 | 2 | 4 | 3 | 3 | 3 | 4 | 7 | 2 | 5 | 5 |
| 55-59 | 6 | 6 | | | 4 | 4 | 8 | 4 | 2 | 4 | 3 | 6 | 5 | 2 | 2 | 1 | 3 | 2 |
| 50-54 | 3 | 1 | | | 3 | | 3 | 1 | 2 | 1 | | 3 | 4 | 2 | 3 | 2 | 3 | 2 |
| 45-49 | 2 | | | | 1 | | 2 | | 3 | 2 | 1 | 1 | | | 1 | | 1 | |
| 40-44 | | | | | 2 | | 1 | | | 1 | | 1 | 1 | 2 | | | 2 | |
| 35-39 | | 1 | | | 1 | | | | | | | | 1 | | | 1 | | |
| 30-34 | | | | | | | | | | | | | | | | | | |
| 25-29 | | | | | | | | | | | | | | | | 1 | | |
| Total | 19 | 18 | | | 23 | 14 | 22 | 16 | 18 | 18 | 7 | 19 | 15 | 15 | 17 | 13 | 18 | 13 |

| 44 | | 45 | | 46 | | 47 | | 51 | | 54 | | 55 | | 57 | | 58 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | Totals |
| | 1 | | | 1 | | | | | | | | | | | | | | 85-89— 2 |
| | | | 1 | | | | | | | 1 | | 1 | | | | | | 80-84— 4 |
| 2 | 4 | 1 | 1 | 1 | 1 | 1 | 4 | 1 | 1 | 1 | 1 | 1 | 2 | | 1 | | | 75-79— 31 |
| 4 | 3 | 2 | 5 | 3 | 6 | 2 | 5 | 1 | 1 | 2 | 2 | 4 | 3 | 1 | 3 | 1 | 3 | 70-74— 77 |
| 3 | 2 | 2 | 2 | 2 | 4 | 4 | 3 | 3 | 2 | 3 | 4 | 7 | 1 | 3 | 6 | 2 | 4 | 65-69— 97 |
| 1 | 2 | 4 | 4 | 4 | 4 | 6 | 1 | 5 | 3 | 6 | 1 | 2 | 1 | 8 | 1 | 5 | 2 | 60-64—125 |
| 3 | 2 | 2 | 1 | 3 | 1 | | 2 | 3 | 4 | 4 | 2 | | 4 | 2 | 2 | 4 | 1 | 55-59—102 |
| 2 | 2 | 1 | 1 | 5 | | 1 | | 2 | 1 | 2 | | 2 | | 3 | 3 | 3 | 3 | 50-54— 63 |
| 1 | | | | | | | | 1 | 3 | | | | | | | 4 | | 45-49— 24 |
| | | | | | | | | | | | | | | | | 1 | | 40-44— 10 |
| | | | | | | | | | | | | | | | | | | 35-39— 4 |
| | 1 | | | | | | | | | | | | | | | | | 30-34— 1 |
| | | | | | | | | | | | | | | | | | | 25-29— 1 |
| 16 | 16 | 12 | 15 | 19 | 16 | 14 | 15 | 16 | 16 | 19 | 10 | 17 | 11 | 17 | 16 | 20 | 13 | 543 |

Table 3-3

California Achievement Test -- First Battery
Range of SD's and Average SD's for Classrooms in the Two Treatments

| | Arithmetic | | | Language | | | Reading | | |
|---|---|---|---|---|---|---|---|---|---|
| | Range | $\overline{SD}$ | Percent Reduced | Range | $\overline{SD}$ | Percent Reduced | Range | $\overline{SD}$ | Percent Reduced |
| Ability Grouped 4th Grades | .32 - .58 | .46[a] | 9 | .48 - .82 | .64 | 33 | .56 - 1.07 | .74 | 23 |
| Random Grouped 4th Grades | .43 - .58 | .50 | | .70 - 1.02 | .87 | | .72 - 1.06 | .91 | |
| Ability Grouped 6th Grades | .40 - .66 | .52 | 19 | .39 - .87 | .67 | 28 | .69 - 1.04 | .86 | 43 |
| Random Grouped 6th Grades | .42 - .81 | .62 | | .66 - 1.13 | .86 | | .84 - 1.46 | 1.23 | |
| Ability Grouped 7th Grades | .44 - .90 | .75 | 12 | .74 - 1.46 | 1.07 | 8 | .81 - 1.28 | 1.01 | 13 |
| Random Grouped 7th Grades | .63 - 1.17 | .92 | | .73 - 1.53 | 1.16 | | .83 - 1.40 | 1.14 | |
| Ability Grouped 8th Grades | .55 - 1.20 | .96 | 22 | .80 - 1.01 | .94 | 42 | .74 - 1.39 | .97 | 44 |
| Random Grouped 8th Grades | .71 - 1.39 | 1.17 | | .78 - 1.68 | 1.34 | | .87 - 1.82 | 1.40 | |
| Ability Grouped 9th Grades | .84 - 1.09 | 1.06 | 36 | .67 - 1.01 | .87 | 49 | .55 - 1.00 | .85 | 43 |
| Random Grouped 9th Grades | .99 - 1.96 | 1.44 | | .96 - 1.85 | 1.30 | | .92 - 1.68 | 1.22 | |

[a] All SD's expressed in Grade Placement units, i.e., .5 indicates one-half year achievement.

Many of the ability grouping studies carried out in the 1920's and 30's, such as Worlton's (1928) study, went to great lengths to maintain identical curricula for pupils at all ability levels in ability grouped classes. In other words, every effort was made to avoid any adjustments in the curriculum that may have better fit it to the superior or slow pupil. Such an approach, of course, overlooks the main purpose of ability grouping which is to reduce the heterogeneity of students within a given class so as to make it easier for the teacher to adjust for individual differences. Thus, in striving for better control, these studies eliminated the principal element they should have studied. An enrichment program such as that employed in District R attempts the same sort of curriculum adjustment and, for that matter, in this research the two districts sometimes employed identical materials for making the adjustment for different ability levels. These adjustments in District R, however, were made in classes in which no attempt has been made to reduce the heterogeneity of the group.

In order to evaluate the results of an ability grouping or an enrichment program, it is necessary that the reader have information concerning the curricular adjustments made in these programs to adapt the curriculum to pupils at different levels of ability. Some of the previously carried out studies have made considerable curricular adjustments in the ability grouped classes but no adjustments in the random grouped classrooms. This also appears to be a questionable approach as, in effect, it involves comparing "something" with "nothing," i.e., a definite program vs. no program to adjust the curriculum to pupil ability. In the research reported herein, both of the districts involved had positive programs and were closely comparable in the resources available for implementing these programs. We will now discuss the curriculum adjustments made in both District A and District R so that the reader may appraise for himself the effectiveness with which the two districts carried out the programs to which they were committed, and may relate these curriculum adjustments to differences in achievement and other variables that will be discussed later.

## Adapting the Curriculum

In general, the goal of the District A program at both elementary and secondary level was to adapt curricular materials to the different ability levels and concurrently adjust the rate of presentation to the level of the child. Thus, at a given grade level the slow group would use curricular materials appropriate to their ability and would be working on learning objectives that had probably been attained by the average groups earlier in the year and by the superior groups during the preceeding year. Adjustment of rate rather than depth to provide for individual differences is characteristic of acceleration programs as opposed to enrichment programs. The curriculum that progresses at different rates for different pupils requires a careful integration between the elementary, junior high school and senior high school programs. Otherwise the bright child will find himself repeating in the seventh grade material he learned in the fifth grade and the slow child will find himself expected to have knowledge and skills that he has never been taught. Perhaps the main advantage of the accelerated program is that it permits the superior student to move to higher levels than is possible in an enrichment program. For example, the pupil who finishes the usual sequence of general science courses a year earlier than usual can take one or two more of the advanced electives in the physical sciences or natural sciences than he could otherwise fit into his high school course of study.

In contrast to District A where the main goal was to adjust rate of learning to individual differences, District R employed enrichment in heterogeneous classrooms to adjust the depth of learning to individual differences. The superior pupil in District R would therefore not take specific courses, such as algebra, any earlier in his program than would other pupils. If successful, however, such an enrichment program should develop a fuller understanding of the work covered by the superior pupil. The enrichment program also provides the bright pupil with more time for independent study and thought than would be available to him in an accelerated program. Whether such pupils generally profit from this free time or instead become lazy and develop poor study habits and attitudes is one of the questions with which the Utah study was concerned.

One would expect the problem of curriculum adjustment to be somewhat more difficult in District R because the typical classroom contained pupils who were more variable than in District A, thus requiring the District R teacher to adjust the materials to a wider range of individual differences.

Let us now discuss some of the specific adjustments made by the two districts in order to provide the reader with greater insight into the setting in which this researc was conducted and the success of the two systems in carrying out adjustments dictatec by their respective policies of enrichment and acceleration.

## Curriculum Adjustments in District A Elementary Classes

No regular in-service training program was carried out for preparing Di[illegible] A teachers in the elementary grades to handle classes of either superior or low ability pupils. However, supervisors worked with the teachers and held study groups and conferences during the course of the study to help teachers improve their adaptation of the curriculum to these groups.

### Reading

Adjustments in the reading program included adopting a series of text books which were appropriate in terms of reading level for the pupils in the superior, average, and slow groups at each grade level. The regular series of textbooks made available in the slow classes went down to grade level 3.1 and a series of textbooks moving up from that level were available. When a child was encountered in any class who required materials not covered in the range of materials made available to classrooms at his grade and ability level, additional materials were obtained from the supervisor on an individual basis. In the ability grouped classes reading was carried out using both intraclass grouping and the whole class instruction.

### Arithmetic

The arithmetic program was also adapted to the three ability levels. This adaptation involved the use of simpler materials and a slower rate of progress at the slow ability level and accelerated progress through more complex materials at the superior level. For example, the grade four text for slow pupils used during most of the program was The New Learning Numbers, from the Winston Arithmetic Text Book Series The extended program text published by the Winston Arithmetic Series, Reaching Number Goals, was used with the average or regular groups. The superior group also used *Reaching Number Goals* but this text was completed in less than the full year, leaving time for half of the fifth grade extended program textbook, *Using Number Ideas* to be covered. In the area of arithmetic relatively little supplemental material was made available to the classroom teacher as the primary goal of the program was completion of the regular sequence in arithmetic at either a faster or slower rate for pupils of different abilities.

### Science

In the elementary science program in District A slow pupils used a text book that was designed for the grade level previous to the one in which the pupils were enrolled. Regular and superior pupils used the regular text book designed for the grade level. Supplemental materials for the slow classes included the basic science education series published by Rowe-Peterson for the primary grades. The superior pupils had available the basic science education series published by Rowe-Peterson for the intermediate grades and junior high school grades. Regular pupils used the intermediate grade level of the Rowe-Peterson Series. Superior pupils also had available the Science Research Associates' Science Series for the junior high school level. Mobile science laboratories were available in all of the District A elementary schools to provide additional supplemental material that could be adapted to the pupils of different abilities. Science was taught in the District A elementary schools using predominantly the whole class instructional approach, with pupils carrying out individual projects.

### Social Studies

The adaptation of the social studies curriculum for superior pupils was made primarily through the use of supplemental materials. Very little adjustment was made in the curriculum for slow pupils in social studies. This is the subject area where the least provisions were made for adjustment of the curriculum in the District A elementary schools. Adjustments in the social studies curriculum were, therefore, left largely to the teachers who were able to adjust the level of projects and other activities to some extent. The fact that satisfactory text book materials aimed at different ability levels were not available, however, limited the curriculum adjustment that could be made.

## Curriculum Adjustments in District R Elementary Classes

Classroom libraries were provided in all District R elementary classrooms. These varied in size but were generally quite extensive, and contained enrichment materials aimed at adapting the curriculum. Most of the units used by elementary school teachers in District R suggested materials and methods for adjusting the curriculum for different ability levels within the classroom.

### Reading

Perhaps the most extensive adjustments made in the District R elementary classes were in the teaching of reading. Reading materials over a wide range of levels were available in each classroom. Intra-class grouping was used for reading instruction. The district also had a somewhat more extensive remedial reading program than was being carried out in District A during the years of this research project.

### Arithmetic

Fewer materials for adjusting the District R curriculum for the different ability levels were available in arithmetic than in the other subject areas considered in this research. Arithmetic teaching was carried out on a whole class basis and the time available to the teacher for working with individual children was limited.

### Science

The science curriculum in District R calls for whole class instruction coupled with individual projects. One or two large portable science laboratories (depending upon the school size) were available in each District R elementary school. Extensive enrichment materials were available in each classroom for adapting the curriculum to both slow and superior pupils although more materials were available for superior pupils than for slow pupils.

### Social Studies

The social studies curriculum at each grade level in District R employed a single text book for all pupils with a moderate amount of supplemental materials in the classroom libraries. The program followed closely the State of Utah social studies guide During the final two years of the study, the National Council for Social Studies Guide was also used.

## CURRICULUM ADJUSTMENTS IN THE JUNIOR HIGH SCHOOLS

### District A Adjustments

In District A, adjustments were made to the science and mathematics curricula to better fit the courses to the ability level of the pupils. These adjustments were made in both content and rate of progress.

### Mathematics

In junior high school mathematics, the superior pupils completed seventh grade mathematics which they had started in the sixth grade and covered the usual eighth grade mathematics courses as well by the end of the seventh grade. These pupils then took Algebra I during the eighth grade and Algebra II in the ninth grade, thus completing four years of mathematics during the three year junior high school program.

Average pupils in District A followed the usual mathematics program carried out in most junior high schools, taking seventh grade arithmetic, eighth grade arithmetic and either Algebra I or general mathematics. This program, excepting the option at grade nine, is very similar to the District R mathematics program in both rate and content.

Slow pupils were given a simplified seventh grade arithmetic course during their first junior high school year. This simplified course was continued through the eighth grade and used many supplemental materials such as the Continental Press Materials which review elementary mathematical concepts down to about the fourth grade level. In the ninth grade, most of the slow pupils in District A took general mathematics, using text books specifically developed for slow learning pupils.

## Science

During this research project, District A used the Holt series of text books in the science curriculum. Superior junior high school pupils took the usual eighth grade course in the seventh grade. The ninth grade course provided by this text book series was taken during the eighth grade and during the ninth grade a course in modern physical science was taken which dealt primarily with chemistry and physics and was similar to the usual high school course in general science. A considerable quantity of supplemental material including high school text books and books to build depth in specific sciences were also used in the superior science classes.

Average pupils in District A studied Book I of the Holt Science Series during the seventh grade, Book II, during the eighth grade and Book III during the ninth grade, thus following the usual sequence at the same rate employed in the more heterogeneous classes in District R.

Slow pupils in District A progressed through part of the regular seventh grade text from the Holt Series and also used a large amount of supplemental material such as the Rowe-Peterson Science Pamphlets which were designed for the upper elementary grades. During the eighth grade, slow pupils continued to work with Book I of the Holt Science Series which they finished by the end of the eighth grade. Again, a large amount of supplemental material designed for slow learning pupils were used along with the textbook at this level. At the ninth grade level, slow pupils in District A proceeded through part of Book II in the Holt Science Series, again using a large range of supplemental materials designed for slow-learning pupils.

In summary, it may be seen that District A adjustments for the different ability groups in mathematics were achieved primarily through adjusting the rate at which pupils covered the mathematics curriculum. Superior pupils covered about four years of work during the three year junior high school period, average pupils covered the usual three year sequence and slow pupils covered what would be two years of work for the average group.

In making adjustments to the science curriculum, District A used both enrichment materials and adjustments in the rate at which these materials were covered. Again, the superior pupils covered work equivalent to that usually covered in four years by average pupils and worked with advanced enrichment materials designed to build greater depth. Average pupils in science followed the usual junior high school science sequence while slow pupils covered the equivalent of two years of the average science program during their three years and worked extensively with supplemental materials designed for slow learning pupils.

## District R Adjustments

## Mathematics

It will be recalled that District R provided curriculum adjustment within the heterogeneous classroom through the use of enrichment. At the junior high school level in mathematics, the Rowe-Peterson Enrichment Series was available for both seventh and eighth grade superior pupils. A set of these materials was available in every mathematics classroom. Additional enrichment materials designed for adjusting the mathematics curriculum to both superior and slow pupils were also available in the central library in each District R junior high school. Curriculum guides in mathematics, however, did not outline special provisions to be made for adapting the curriculum to slow and superior pupils. Some in-service training, in adapting the curriculum through the use of enrichment materials, was given District R teachers.

### Science

In science, a very large selection of enrichment materials were available at both the junior high school and senior high school levels, these materials being purchased for the most part through an NDEA grant. Science teachers carried out surveys to identify enrichment materials to adjust the curriculum and were thus familiar with the materials available. A separate curriculum guide was available in science for adjusting the junior high school curriculum for superior pupils but such a guide was not available for slow pupils.

Thus, we see that a considerable amount of enrichment material was available at the junior high school level in District R mathematics and science classes. Considerably more enrichment materials were available for adapting the curriculum to superior pupils as contrasted with slow pupils. At the junior high school level many of the materials used in District R for adapting the curriculum for heterogeneous classes through enrichment were also used for the same purpose in the ability grouped classes of District A. Both districts use the Rowe-Peterson Enrichment Series, for example. More materials were available in District A for making curriculum adjustments for slow pupils in both mathematics and science. Adjustments in District A junior high school mathematics, however, were largely in terms of changing the rate of instruction although supplemental materials were used to a considerable degree for slow pupils at this level. Science instruction in District A makes extensive use of supplementary materials, in addition to adjusting the rate of instruction.

## CURRICULUM ADJUSTMENTS IN THE HIGH SCHOOLS

At the high school level, regular and superior pupils in District A were grouped in Algebra II and geometry. As no slow pupils enrolled in these courses, the slow groups were not present. Near the end of this study, when District A had followed its grouping system long enough for superior pupils to have completed Algebra II in the ninth grade, the number of tenth grade superior pupils grouped in Algebra II was largely limited to transfer students and those who had not taken ninth grade mathematics. It was found that no grouping was necessary in the more advanced mathematics courses such as solid geometry and trigonometry as superior pupils were the only ones taking these courses with very few exceptions. Continuing the pattern of adjustment through acceleration, followed at all levels in District A, superior pupils in mathematics could go on to analytical geometry and differential calculus in the final high school year. This option, however, was not available until near the end of the study as it required three years in the ability grouping program for superior pupils to reach a point in the accelerated curriculum where they were ready for these college level courses.

Biology was the course taken to meet the high school science requirements by a large majority of pupils in District A. Biology courses were grouped into three ability levels. Chemistry was grouped into two levels with no slow pupils taking the course. The other elective science courses were not grouped because of the relatively small number of pupils taking these courses and the fact that most pupils taking these electives belonged to the superior group.

### Limitations at the Secondary Level

The reader will note from the data given on the treatments and curriculum adjustments, that as pupils moved from elementary to junior high and finally to high school, it became increasingly difficult to control some of the variables pertinent to the study. At the elementary school level in which both districts followed the Utah State Study Guides rather closely and in which self-contained classrooms were used, a higher degree of control was possible and the data obtained appear to be most reliable. At the junior high school level, both districts again followed closely the Utah State Curriculum Guides in science and mathematics and both, of course, met the minimum state requirements in these subjects. At the ninth grade level, however, control became more difficult. Although the two districts had the same requirements, District A offered both algebra and general mathematics and strongly encouraged pupils to take one of these two courses. At the high school level, controls became even more difficult because only one year of mathematics and science was required. The electives taken by pupils thus became difficult to control and differences between the ability grouped and random grouped approaches became increasingly smaller through the high school years. Ability plays a large role in all secondary schools in determining which students will take the more advanced electives in science and mathematics. Thus, the findings of this study at the elementary level, it is felt, can be accepted with considerable confidence. The results for the junior high school level, although subject to the limitations mentioned above, are still felt to be reasonably valid and well founded and should have many implications for teachers and administrators. At the high school level, however, the findings should be regarded at best as tentative and suggestive. Results at the early high school level are somewhat more meaningful as fewer electives were available during the tenth grade and furthermore, tenth grade results can be regarded to a considerable degree as reflecting the effects of the earlier work at the junior high school level. At the eleventh and twelfth grades, however, the wide range of electives available and the tendency for ability grouping to occur naturally in both districts because of the restricted ability range of pupils choosing mathematics and science electives lead the principal investigator to suggest that the reader interpret findings at this level with caution.

## SUMMARY OF SAMPLING AND TREATMENT DIFFERENCES

The larger and, in most respects, more comparable samples employed in this study were those initially selected at the elementary school level--Samples IV and VI. Comparability of elementary schools was somewhat less difficult to establish as these schools serve smaller and somewhat more homogeneous areas in terms of socioeconomic status. During their elementary school years, these pupils were also exposed to more comparable curricula in the two districts. Also, as both districts used self-contained classrooms, the treatments, especially in District A, influenced a greater portion of the pupil's day than was the case for secondary school samples. In other words, the District A pupil was in an ability grouped class during the entire day and the District R pupil was in a heterogeneous or random grouped class for the entire day at the elementary school level.

The data obtained at the junior high school level provided a somewhat less satisfactory comparison of the two grouping treatments for several reasons. First, it was not possible to involve junior high schools that were as closely comparable in socioeconomic status. Also, and especially at the ninth grade level where electives began to be introduced, the junior high school curricula began to differ more sharply between the two districts, thus introducing differences that make it more difficult to obtain clear-cut results. Finally, District A employed ability grouping in mathematics and science only at the junior high school level, thus limiting the extent to which the different grouping treatments affected the pupil during his school day.

As Samples VII, VIII, and IX moved up into the high school level, treatment comparisons became even more difficult largely because of the wide range of electives in mathematics and science. In view of these limitations relatively fewer pupils were selected in Samples VIII and IX and fewer measures were administered to these samples. Thus, in interpreting the results of this research it is believed that within the confines of the ability grouping and random grouping treatments employed in these two school districts, the elementary school data and the data gathered during the first two years of junior high school are much more meaningful and can be accepted with greater confidence than those dealing with data collected during the ninth grade and the high school years.

## REFERENCES

French, J. W. Evidence from school records on the effectiveness of ability grouping. J. Educ. Res., 1960, 54, 83-91.

Goldberg, Miriam L., Justman, J. A., Passow, A. H., & Hage, G. The effects of ability grouping. Unpublished manuscript, New York: Teachers College, Columbia Univer., 1961.

Worlton, J. T. The effect of homogeneous classification of scholastic achievement of bright pupils. Elem. School J., 1928, 28, 336-345.

# Chapter 4

# ACHIEVEMENT

## INTRODUCTION

The goal of this chapter is to explore the relative effectiveness of ability grouping and random grouping in terms of achievement gains made by pupils under these two systems. We will examine the rather extensive previous research that has been carried out in this area, and will then add the findings of this project to the previous knowledge that has been assembled. In the Utah Study, achievement data were collected for all subjects at the beginning of the study, and on at least three additional occasions during the subsequent four years. Achievement data were collected on at least one of the five samples at every grade level from fourth to twelfth inclusive, and for some grade levels, data were collected on as many as four different samples.

## WORKING HYPOTHESES

A set of statistical hypotheses in the null form were developed as part of the original research plan and may be found in Appendix C. The working hypotheses on achievement which formed the basis for designing this phase of the project were founded primarily upon psychological rationales developed by the author or put forth in the writings of other psychologists and educators. Previous research was also used as a basis for developing working hypotheses, but because of the design limitations characteristic of most previous work, one finds this work a poor foundation upon which to build hypotheses. The major working hypotheses relative to achievement were:

1. Superior pupils will achieve more in ability grouped classrooms in which the curriculum is adjusted primarily through changes in learning rate, than they will in random grouped classrooms in which enrichment is the primary means of curriculum adjustment.

This hypothesis has considerable basis in the findings of previous research. A number of studies such as the work of Provust (1960), West and Sievers (1960), and Herr (1937) for example, have found greater achievement gains for superior pupils in ability grouped classrooms than for superior pupils in heterogeneous classrooms.

Educators have also made certain assumptions about the ability grouping situation which would tend to lead one to expect superior pupils to make greater gains under this treatment. It is often maintained, for example, that the superior pupil is seldom challenged in the random grouped classroom and as a result tends to work at a level far below his potential. In contrast, the ability grouped classroom is looked upon as providing a situation in which the superior pupil is faced with an accelerated program and more intense competition from his peers and these factors tend to motivate him to higher achievement.

2. With respect to the achievement of average pupils in ability grouped and random grouped classrooms, it was our working hypothesis that no significant differences would emerge.

Both the findings of previous research and the rationale underlying ability grouping tend to lead to this hypothesis. A number of studies such as Worlton (1928), Provust (1960), Billett (1928), for example, have obtained greater differences in achievement for pupils in high or low ability extremes than were found for pupils around the center of the ability distribution. Research results comparing ability grouped and random grouped average pupils seem to be about equally divided in favoring one system or the other, thus suggesting no consistent trend.

A frequent criticism of heterogeneous grouping is that the curriculum is geared to the average pupil. If this is the case, the average pupil would have less to gain by moving to an ability grouped situation because the random grouped situation is already adapted to his level. Some educators have maintained that the average pupil suffers in the ability grouped situation as he is deprived of the assistance and contributions of superior pupils in the classroom situation. Some research, however, (Mann, 1957) suggests that very little real contact takes place between pupils of different ability levels, thus minimizing the possible gains that could occur through interaction between pupils of different abilities.

3. It was hypothesized that slow pupils would make greater achievement gains in the random grouped classroom situation. This working hypothesis was somewhat harder to establish than those proviously discussed, because of a number of apparently conflicting pressures and motivations that appear to be operating upon the slow pupil in the two grouping situations. Some of these pressures would appear to raise the achievement of the slow pupil in the ability grouped classroom, while others would appear to lower it. For example, the fact that the rate of learning in ability grouped classrooms can better be adjusted to the capacities of the slow learning pupils would lead one to expect these pupils to achieve more in the ability grouped situations. On the other hand, the stigma that most educators believe to be associated with membership in a slow group might well develop unfavorable attitudes, reduce motivation, and lower the self-concept and aspiration level of slow pupils in ability grouped classes, which would be expected to have a detrimental effect on their achievement.

In random grouped classrooms it has been argued that the class goals are so far beyond the ability of the slow pupil that he becomes discouraged or disinterested. On the other hand, it may be contended that the random grouped classroom in effect provides an accelerated situation for the slow pupil, the pressures of which might lead him to greater achievement, just as similar pressures in ability grouped classrooms are thought to motivate the bright pupil. The reader will note in the review of the literature that research dealing with the achievement of ability grouped slow pupils reports somewhat conflicting results. Some studies have found no differences between the achievement of slow pupils in ability grouped and random grouped situations (Moyer, 1924; Hartill, 1936; Provust, 1960). Others have shown relatively greater gains for slow pupils in ability grouped classes (Billett, 1928; Rankin, 1931) while a study by Cook (1924) found that ability grouped slow pupils achieved more in some subjects while random grouped slow pupils achieved more in others.

## PREVIOUS RESEARCH RELATED TO ACHIEVEMENT AND GROUPING

### Introduction

Over the years, perhaps the major appeal of ability grouping to the educator has been its promise to provide a classroom situation in which better adjustments of the curriculum for individual differences might be possible. The goal of such curricular adjustments generally is to bring about higher achievement by providing more adequately for individual differences among pupils. In view of this appeal, it is not surprising to discover upon reviewing the research literature related to ability grouping that the vast majority of studies in this area are concerned with comparisons of achievement gains in ability grouped and random grouped classrooms.

A number of reviews of research in this field have been published. The reader is recommended to Ekstrom (1959) who has done a thorough job or reporting and critically evaluating the research on ability grouping up to the date of her review. Rather than duplicate this and other excellent reviews that have been published, we will in this chapter review in detail only a few outstanding studies, many of which have been completed since Ekstrom's review was published, This rather detailed coverage of what appear to be the better studies in this area has been supplemented by two tables which give the essential information on a large number of additional studies that may interest the reader. Table 4-1 has been compiled for studies that have been carried out at the elementary school level, while Table 4-2 deals with studies at the secondary school level. Even these tables are somewhat selective as they do not include many of the weaker experimental studies nor any of those publications which present opinions or personal experience rather than research evidence.

In our discussion of those studies that we believe present the best evidence related to ability grouping, we have pointed out a number of weaknesses and limitations. These criticisms are not made with the intention of minimizing the importance of the knowledge that these studies have given us, but instead attempt to provide the reader who is unfamiliar with this research area with some insights that will help him to place the various studies in their proper perspective.

## Research at the Elementary Level

Although completed more than 30 years ago, Hartill's research (1936) is probably the best designed work published to date on the relative merits of homogeneous and heterogeneous grouping at the elementary school level. The subjects in Hartill's study included 1374 pupils enrolled in grades 5A, 5B, and 6A. These pupils were tested in December, 1931, using the Stanford Achievement Tests, Form V. On the basis of educational age, one-half of the subjects were grouped homogeneously into three ability levels while the other half were assigned to heterogeneous groups. Pupils in the homogeneous groups received a differentiated course of study aimed at their ability level, while those in the heterogeneous groups received an undifferentiated course of study. In June of 1932, after the subjects had been exposed to one-half year in the homogeneous or heterogeneous classes, they were administered the Stanford Achievement Tests, Form W.

The following September, when these pupils returned to school, those who had been in homogeneous groups during the previous school term were placed in heterogeneous groups, and vise versa. Teachers were advanced also so that in a given school the teachers who had taught 5A children who were grouped homogeneously would teach the same children in grade 5B when they were placed in heterogeneous groups. At the end of the fall term in December of 1932, all pupils were again administered the Stanford Achievement Test, Form V.

An overall analysis of the achievement results for pupils when in homogeneous and heterogeneous groups showed no significant differences between the overall gains made

Table 4-1

Ability Grouping Studies Concerned with Achievement at the Elementary School Level

| Study | Grade Levels | N | Length of Study | Grouping Levels | Grouping Criteria | School Subjects Measured | Results | Remarks |
|---|---|---|---|---|---|---|---|---|
| Jones McCall (1926)[a] | 3-7 | 67 Pairs | 2 years | Superior Only | I.Q. | Reading<br>Arithmetic<br>Spelling | Slight advantage for AG sup. on groups, not sig. | AG pupils elected to enter accelerated classes, thus biasing sample. Matched pairs used. Curriculum differentiated for AG pupils. |
| Worlton (1926) | 4-7 | 3700 AG & RG | ½ year | 3 | I.Q. | Reading<br>Arithmetic<br>Spelling | AG classes made greater gains in composite ach. score (no mention of stat. sig.) | Steps taken to avoid differentiation of curriculum for dif. ability levels. |
| Baird (1927)[a] | 4-8 | 325 in AG + 10 classes RG | | 3 | Ach. Test | Gen. Ach. | 8 of 10 control classes achieved more (no mention of stat. sig.) | |
| Walton (1928) | 4-7 | 1466 | | I.Q. Levels | Teacher Judgment & I.Q. | Reading<br>Spelling<br>Arithmetic | Slight advantage for AG sup. groups on composite ach. score (no mention of stat. sig.) | 1466 pupils above 90 I.Q. were tested. Pupils were classified into 6 catagories each with a range of 10 I.Q. points. Curriculum differentiation implied but type of adjustments not described. |
| Bonar (1929) | 1 | 96 | 1 year | 2 | I.Q. & Teacher Judgment | Reading | Results on objective ach. test favored RG. No data on stat. sig. | The grouping system used did not result in really dif. ability groups.The median I.Q. of the high group was 104, the low group 98 & the het. group 99. |
| Dvorak, Rae (1929) | 1 | 20 Pairs | 1 year | 2 | I.Q. | Reading<br>Spelling | No clear-cut dif. bet. AG & RG. Two stat. sig. dif. in reading scores found favored AG. | The grouping system used could not be expected to reduce the variability of the AG class by more than a few percent. |
| Gray Hollingsworth (1931) | Age 7-9 | 56 exp.<br>36 control | 3 years | Superior Only | I.Q.(S.B.) Above 130 | Reading<br>Arithmetic<br>Nature<br>History<br>Language | No stat. sig. dif. found in regular school subjects. "Enrichment" subjects studied by AG pupils not measured. | AG pupils did extra work in some subjects and studied some topics not covered by RG pupils |
| Barthelmess Boyer (1932) | 4-5 | 297 Pairs and 268 Pairs | 1 year | 3 | I.Q. | English<br>Reading<br>Geography<br>Arithmetic | Total AG sample was sig. higher on total ach. for both replications. AG pupils at all 3 ability levels made greater gains but sig. level of these gains not reported. | Study involved 2 replications. Used matched pairs which resulted in loss of up to 53% of AG pupils & 77% of RG pupils. |

| Study | Grade Levels | N | Length of Study | Grouping Levels | Grouping Criteria | School Subjects Measured | Results | Remarks |
|---|---|---|---|---|---|---|---|---|
| Hartill (1936) | 5-6 | 1374 | 1 year | 3 | Ach. test | Gen. Ach. | Overall comparisons bet. AG & RG pupils showed no stat. sig. dif. Sup. RG pupils & av. AG pupils made greater gains. No dif. bet. slow groups. | Well controlled study using counterbalanced design, to control group dif. Differentiated course of study used with AG sup. pupils being given enrichment program rather than acceleration. |
| Rankin, et. al. (1936) | 3-6 | About 500 in each plan | 2 years | 3 | I.Q. | Reading<br>Arithmetic | On composite score of 4 reading tests, vertical plan best at all ability levels. Detroit plan higher than RG for composite & bright pupils but these 2 plans about same for av. & slow pupils. In arithmetic Detroit plan best, vertical plan next and RG lowest for all ability levels. Dif. greatest for sup., smallest for slow groups. | Two dif. AG plans compared with RG plan. Vertical plan cut across grade levels to reform AG; Detroit plan grouped within grade levels. |
| Kvaraceus Wiles (1938) | 3 | 75 | 1 year | 3 | Ach. Test | Reading<br>English<br>Arithmetic | "Better than av. pupil growth," no mention of stat. sig. | Pupils grouped separately in reading, English, & Arithmetic, & grouped het. rest of day. Major effort made to differentiate curriculum. No control group. |
| Russell (1946) | 4-6 | 278 exp.<br>248 control | 2 years | 3 | Reading Ach. | Reading | No stat. sig. dif. in gains for either the entire groups or for pupils at dif. ability levels. | Involves AG in reading only with pupils in het. groups for rest of school day. A well designed study. |
| Edmiston et. al. (1949 | 5-6 | 462 | 1 year | See Remarks | I.Q. Range | Reading | Wide range group made sig. greater reading gains. | Compared "wide range" groups (I.Q. Range 41) & "narrow range" groups (I.Q. range 29). Curriculum not differentiated. |
| Husen Svensson (1960) | 5-6 | 2755 | 2 years | 2 | School Marks Based on Standard Test | Reading<br>Grammar<br>Spelling<br>Arithmetic | As analysis was broken down by treatment, social class, & ability level, a brief summary of results is not possible. Few sig. dif. found bet. treatments for pupils of comparable ability. | This study was carried out in Sweden. |
| Provust (1960) | 4-6 | 254 | 1 year | 3 | I.Q. | Arithmetic | Overall comparison gave sig. greater gain for AG. Analysis by level showed AG sup. made greater gains. No dif. bet. AG & RG av. or slow pupils. | Experimental group in AG class for arithmetic only' in RG class for rest of day. Study used matched pairs. |

| Study | Grade Levels | N | Length of Study | Grouping Levels | Grouping Criteria | School Subjects Measured | Results | Remarks |
|---|---|---|---|---|---|---|---|---|
| West Sievers (1960) | 5-6 | 16 exp.<br>16 control |  | Superior Only | Teacher Judgment & Ability Test | Language Art & Arithmetic | Mean gains of 2 to 5 months for AG over RG pupils. | Matched pairs design. Use of teacher judgment probably introduced bias favoring experimental group. Experimental group was in special AG class half day & RG class reaminder of day. |
| Daniels (1961) | Junior School 1-4 yrs. | About 500 cases. Varies from year to year. | 4 years | 3 | I.Q. & previous Grades. | English<br>Reading<br>Arithmetic | 16 dif. bet. means for one pair of schools, 4 for another pair of schools. All 20 reported sig. favoring RG schools. No breakdown of data by ability level. | No data given on curriculum differentiation in AG schools. Groups matched for mean & S.D. of I.Q. Author apparently used one-tailed test of sig. as CR's as low as 1.75 are reported sig. at .05 level. |
| Goldberg et. al. (1961) | 4 | 2219 |  | 5<br>See Remarks | I.Q. | Reading<br>Arithmetic<br>Language Arts, Work Study Skills<br>Science<br>Soc. Stud. | Of 105 ach. comparisons, there were 11 that were stat. sig. 10 of these favored "broad range" patterns i.e. het. grouping. | 5 ability levels combined into classes to provide 15 patterns differing in ability levels & degree of hom. Curriculum not differentiated. |
| Koontz (1961) | 4 | 191 to 231 for dif. subjects | 1 year | 5 | Arithmetic<br>Reading<br>Language | Arithmetic<br>Reading<br>Language | RG sample achieved sig. more in reading & arithmetic. No sig. dif. in language. No sig. interaction bet. treatment & level. | AG & RG pupils matched on initial ach. Analysis of variance used to analyze results. Hawthorne Effect probably favored AG. |
| Simpson Martinson (1961) | 5-6 | 295 in Grades 5-6 | 1 year | Superior Only | I.Q. (S.B.) Above 130 | Arithmetic<br>Reading<br>Writing<br>Listening<br>Science<br>Social Studies | Both the special enrichment group & special hom. group achieved sig. more than control group on gen. ach. | This study involves a wide variety of treatments for sup. pupils in grades 1-12. All treatments are compared with control groups but treatments are not compared with each other. |

---

[a] Taken from Ekstrom, (1959). Primary source not available.

Abbreviations:
AG = ability grouped
RG = random grouped
stat. = statistical, statistically
sig. = significance, significant
sup. = superior
av. = average
dif. = difference(s), different
hom. = homogeneous
het. = heterogeneous
ach. = achievement
bet. = between
gen. = general

by these pupils under the two conditions. When pupils were subdivided in terms of ability level, however, it was found that the superior pupils made larger gains in the achievement areas tested (except reading) when under the heterogeneous grouping system. Average pupils made larger gains when under the homogeneous grouping system, and slow pupils did equally well under the two systems. It should be noted that superior pupils in the homogeneous groups were given an enrichment program rather than an acceleration program, and were, in fact, held back to a normal degree of growth on the fundamentals so as to permit more time for enrichment. Hartill concluded that the homogeneous plan was superior because bright pupils did nearly as well on the fundamentals and in addition had time to study enrichment materials, and because average pupils gained significantly more, and slow pupils did just as well under this system.

Hartill's research is the only study found in the area of ability grouping in which a counterbalanced research design has been used. This design has the great advantage of cancelling out differences between experimental and control groups because each pupil is measured in both treatments during the two phases of the study. Hartill's design also controlled teacher differences, by having the same teachers move ahead one-half grade so that within a given school, the same teachers would teach both homogeneous and heterogeneous classes.

Another of the early studies which involves large samples, and gets at basic questions with a reasonably sound research design was carried out by Rankin, et. al. (1936). This study involved about 500 pupils in each of three grouping plans, and compared achievement gains on four reading tests and three arithmetic tests over a two-year period. The three grouping plans used were: (1) the "vertical plan" in which pupils were grouped on three ability levels combining as many as three half-grades to make up a class, (2) the "Detroit plan" which also employed three ability levels, but did not cut across grade lines,and (3) the "mass instruction plan" in which no attempt was made to establish homogeneous groups. Curriculum adjustments were made in both of the ability grouping plans although these adjustments were greater in arithmetic under the Detroit plan than under the vertical plan. In the mass instruction plan "... a deliberate attempt was made to teach all pupils in the same way. This was done by using the same materials and the same methods for all and attempting to hold all pupils to the same standards. Standard Detroit materials designed to promote adaptation to individual differences were replaced in these schools by other materials better adapted to mass instruction." (p. 279)

The results of this research are given in terms of gain scores over a two-year period. Gains are given for all pupils on each of the tests and are also presented for bright, average, and slow groups separately. Although the investigators present no data on statistical significance, the information that they provide is sufficient for calculation of t-tests between mean gains. These t-tests were calculated for the Stanford Reading Test, which was weighted about half of the composite reading score and for the Detroit Arithmetic Test, which was weighted about half of the composite arithmetic score. The data reported did not permit calculation of t-tests for mean differences on the composite scores.

The results on the Stanford Reading Test comparing all pupils regardless of ability level showed the vertical plan to be superior to both the Detroit plan and mass instruction plan with the differences significant beyond the .01 level. Differences between the Detroit and mass instruction plan were not statistically significant. When comparing bright pupils only on the Stanford Reading Test, the results again showed that the vertical plan resulted in significantly greater gains than the other two plans, with these differences also significant at the .01 level. Differences between average pupils followed the same pattern with the vertical plan being significantly higher at the .01 level. The pattern was somewhat different, however, for the slow group. Slow pupil gains under the vertical plan were significantly higher at the .05 level than under the Detroit plan, but were not significantly different from the mass instruction plan. Again, the Detroit and mass instruction plans were not significantly different from each other.

It is interesting to note that there were no significant differences between heterogeneous grouping (mass instruction) and homogeneous grouping within the grade level (Detroit plan). Those significant differences found, favored the more extreme system of homogeneous grouping across grade levels that was employed in the vertical plan.

The t-tests calculated by the reviewer in arithmetic were based on the Detroit Arithmetic Test which was weighted about one-half of the composite arithmetic score. Again, data available in the original study did not permit calculation of t-tests for the composite arithmetic score. The data for all pupils combined as well as for the bright, average, and slow pupils separately showed a somewhat similar picture excepting that the Detroit plan yielded the greater gains, usually followed by the vertical plan. The mass instruction plan (heterogeneous grouping) was poorest. The differences in arithmetic, however, were much smaller than those found in reading. None of the differences between the three treatments on arithmetic gains were statistically significant. The largest difference was between the Detroit plan and the mass instruction plan for average pupils and this difference was far short of statistical significance.

Probably the greatest limitation in the findings of this study was brought about by the removal of all enrichment materials, even those normally used by the district, from the heterogeneous classrooms involved in the "mass instruction plan." In taking this step, the investigators created an unusual difference between the treatments in the amount of supplemental materials available, and as a result, the achievement difference obtained are partially attributable to the grouping treatments and partially attributable to the differences in materials available for adapting the curriculum for individual differences. The fact that reading differences were much larger than arithmetic differences suggests that the absence of material for differentiating the curriculum might have had a larger influence on the results than the grouping plan, because the teaching of reading without any materials available for adapting to individual differences presents the teacher with a far more difficult task than teaching arithmetic under these conditions. Further, the superiority of the Detroit plan in arithmetic is explained by the investigators as being probably due to a special individualized program used in this plan which was not used in the vertical plan or mass instruction plan. In spite of this limitation, however, this is one of the best of the comparative studies as it involves reasonably large samples, uses reliable objective measures, breaks down the effect of the treatments for pupils of different ability levels, and avoids the pitfalls of using matched pairs.

Among the more recent studies, the work of Daniels (1961) is one of the most extensive at the elementary level. In this study of four British junior schools (grades 2-5 inc. in our system); two employed three level ability grouping based on I.Q. and two employed random grouping. Although little information is provided on the curriculum adjustment within the ability grouped schools, information provided indicates that the four schools had comparable curricula and that the random grouped schools used whole group instruction rather than enrichment or intraclass grouping. Schools A and B employed random grouping while schools C and D employed ability grouping or "streaming" as the English call it. Schools A and C were found to be comparable on a number of pertinent variables as were schools B and D. Group matching procedures were used to equate the groups for mean and standard deviation in I.Q. This matching resulted in losing relatively few cases. Three achievement areas: reading, English, and arithmetic, were tested at the start of the study and at intervals over a period of four years. Three successive classes entering each junior school were used in the research. For schools A and C, the combined number of cases for different year intervals ranged from 262 to 297. For schools B and D the combined N ranged from 197 to 227.

In comparing mean achievement differences in the three subject areas between school A and school C, Daniels reports all 16 critical ratios between the two treatments at various intervals ranging from a half year to four years to be statistically significant. Taking differences over the longest periods reported, comparisons over a four-year period in reading achievement yielded a critical ratio of only 1.87. Comparisons in English achievement over a three-and-a-half year period yielded a critical ratio of 3.6, while the critical ratio over a three-and-a-half year period in arithmetic was 3.04. All of these differences between means were in favor of school A, the school which employed random grouping. Only four comparisons between schools B and D are reported in the achievement areas but all four of the critical ratios obtained were statistically significant favoring school B, the random grouped school.

I.Q.'s taken over the three-and-a-half year period were also shown to have increased significantly more for both of the random grouped schools. These I.Q. changes were rather small--a superiority of 2.79 points for school A over school C and 2.9 points for school B over school D. All data were analyzed for total groups rather than being broken down by ability levels. Daniels used the rather unwieldy procedure of rematching his groups at various intervals in order to measure gains at different periods over the four years. For example, to measure the significance of gains during the third and fourth years in English, he rematched his groups on the English test given in the second year and then compared the mean English test scores that the groups so matched earned during the fourth year. Analysis of covariance would have been a far simpler and more effective tool for this study.

Daniels' study appears generally well designed and carefully carried out. It is interesting to speculate whether the results, which favor random grouping to a degree not found in any other research represent true differences between the treatments or are due to factors unique in the research situation. If the data had been analyzed separately for pupils at different ability levels, this study would have provided a great deal more useful information.

Another recent study by Goldberg et. al. (1961) is by far the most extensive project to date aimed at evaluation of ability grouping per se, that is, without any attempt at adjustment of the curriculum for groups of different ability levels or degrees of homogeneity. This research was carried out in the New York Public Schools starting in spring of 1956. Subjects were selected from the fourth grade level excluding schools having fewer than five fourth-grade pupils with I.Q.'s over 130. Five ability levels were identified, ranging from group A with I.Q.'s of 130 or higher to group E having I.Q.'s of 99 and below. Classes were organized so as to contain different combinations of the various ability levels. For example, pattern one contained only pupils with the I.Q.'s of 130 and above (Group A), pattern two included pupils from both groups A and B;

pattern three included pupils from groups A, B, and C, and so on. A total of 86 classes were formed which were classified into a total of 15 patterns representing different combinations of pupils at the five I.Q. levels. A wide range of measures were administered to the sample including academic achievement in reading, arithmetic, language arts, work study skills measured by the SRA Achievement Series, academic achievement in science and social studies measured by the Stanford Achievement Tests; interests, attitudes toward self, school and other pupils, and measures of social status and creative expression. Findings of this study in general indicate that when no attempts are made to adjust the curriculum for different ability levels, the process of narrowing the ability range in the classroom is unlikely to raise pupil achievement. In this research, as a matter of fact, pattern five which represented the broadest ability spread among the 15 patterns employed appeared to be most consistently associated with greater academic gains for all pupils. The achievement differences between patterns, however, tended to be small with no one pattern or combination of patterns resulting in better achievement for all pupils in all subject areas.

The findings of this study indicated that achievement was influenced more strongly by the teacher and group differences in individual classrooms than by ability range of the class. Advocates of heterogeneous grouping have often argued that the presence of superior pupils tends to lead to greater achievement of pupils of lower ability. This contention was tested in Goldberg's study and it was found that only in science was the presence of gifted pupils consistently related to greater increments for the other four ability levels. In other school subjects measured, the presence of gifted pupils was not related to greater achievement gains for non-gifted pupils.

A further analysis of the data was accomplished by combining the 15 ability patterns into three broad categories: narrow range, medium range, and broad range classes. An analysis of variance indicated that the ability range in the classroom had a significant bearing on the achievement in social studies, reading comprehension, vocabulary, arithmetic, and total average achievement; while science, language arts, and work study skills were unaffected by the range. In the six academic areas affected, overall comparisons showed greater gains for pupils in the broad range classes as compared with the medium or narrow range patterns. When each ability level was taken separately, the effects of range were inconsistent, although revealing some significant differences. Of the 105 possible comparisons between pupils in the five ability levels, the three range patterns and the seven academic areas only 11 reached significance, and in 10 of these the broad range was superior to either or both of the other two ranges.

This study also explored the question of whether the pupils' position in his group in terms of ability was related to his achievement. This is an important question as one of the commonly used arguments against ability grouping in that it deprives the slow pupil of the chance of learning from more able classmates. The analysis showed that in most subjects the presence of one or two levels of more able classmates and no slower ones did not have any positive effect upon achievement. The investigators concluded that the few observed differences indicating advantages or disadvantages associated with being in a group of higher ability, lower ability, or equal ability pupils were probably due to chance.

The research carried out by Goldberg and her associates is very well defined and involves a thorough and sophistacated analysis that deals with most major questions concerned with classroom level and degree of heterogeneity as it relates to achievement.

When the overall results of the study are considered, one must conclude that variations in the ability level, the ability range, and the position of individual pupils within the ability range in classes where no specific curricular adaptations are made, probably has little effect upon pupil achievement. In other words, ability grouping per se offers no achievement advantages over random of heterogeneous grouping and appeared less effective than broad range grouping under the conditions present in Goldberg's study. This is one of the very few areas related to ability grouping in which scientific evidence from an extensive and well-designed study seems to have presented a clear and reliable answer. An answer to the more fundamental question of whether ability grouping offers achievement advantages when the curriculum has been differentiated is not found in this or other research to date which is concerned with pupils at the elementary level.

## Research at the Secondary Level

Nearly all studies concerned with ability grouping versus random grouping at the secondary level have deficiencies which raise serious questions about the validity of their findings. We will explore some of these deficiencies in the following section and will limit this review to a few of the recent studies that appear to provide reasonably valid research evidence. The reader is referred to Table 4-2 for a brief summary of the major projects concerned with ability grouping at the secondary school level.

Starting in 1956, an experimental four-track curriculum was established in the Washington, D.C., high schools. The following year this program was extended to the eleventh grade and in 1958 to the twelfth grade. When the first class progressing through the entire sequence had graduated, an evaluation of the program was carried out and reported (Hansen, 1959). This four-track program was essentially an ability grouping program designed to reduce the ability range of each classroom and increase efficiency in instruction by differentiating the curriculum. The four tracks included the honors track which was generally restricted to pupils above the 115 I.Q., the college preparatory track, the general track, and the basic track. The honors track was subjected to an extensive comparative evaluation in which seniors from the class of 1958 who were not in the track program were compared with seniors from the class of 1959 who were in the honors track during their high school years. The study involved the top 270 graduates of each year, the groups being selected so as to give an even proportion of boys and girls and very similar composition in terms of I.Q. Both groups had been administered the Iowa Tests of Educational Development during their senior years. The median score for the honors track students of 1959 was five points higher than the comparison group of 1958 in "correctness of expression," three points higher in "quantitative thinking," and six points higher in "science achievement." The two groups were also compared on the College Entrance Board Examination. The honors program group was 37 points higher on their median scholastic aptitude score, and 30 points higher on their median mathematics score. Although the research report does not give levels of statistical significance for these differences, it appears likely that all of the differences reported are significant. This study presents rather sound evidence in favor of a highly differentiated curriculum for superior students. A thorough analysis of the other three tracks as they compare with heterogeneous grouping would add greatly to the value of this study.

Probably the best controlled and most informative study concerned with ability grouping at the secondary level that has been reported to date is the research recently

Table 4-2

Ability Grouping Studies Concerned with Achievement at the Secondary School Level

| Study | Grade Levels | N | Length of Study | Grouping Levels | Grouping Criteria | School Subjects Measured | Results | Remarks |
|---|---|---|---|---|---|---|---|---|
| Theissen (1922) | 7 | 65 to 74 classes | 1 sem. | 6 | I.Q. | Arithmetic<br>Reading<br>Language | Compares AG sections at dif. ability levels rather than comparing AG & RG classes. The poorest of the top I.Q. classes achieved less in both arithmetic and language than the best of the bottom I.Q. classes. | Demonstrates the importance of teacher dif. |
| Cook (1924) | 9-10 | 495 | 1 sem. | 2 | Past Grades | English<br>History<br>Geometry | No dif. bet. AG & RG in English. Sup. AG better in history, poorer in geometry. Slow AG poorer in history, better in geometry. | |
| Moyer (1924) | 9 | | 1 sem. | 3 | I.Q. | Algebra<br>Latin | Sup. AG better in Latin. Av. AG better in both subjects. No dif. bet. slow AG & RG. | No data on stat. sig. |
| Purdom (1924)[a] and (1929) | 9 | | 1 sem. | 3 | I.Q. | English<br>Algebra | No dif. bet. AG & RG. | |
| Woody (1924) | 9 | 51 Matched Pairs | 1 sem. | 3 | I.Q. | English | No dif. bet. AG & RG. Results not broken down by ability level. | Used matched pairs. |
| Martin (1927)[a] | 7 | | 1 yeat | 3 | I.Q. | General Ach. | No sig. dif. bet. AG & RG on general ach. test. | Used matched pairs. Teachers told not to differentiate the curriculum. |
| Van Wagenen (1927)[a] | 7-8 | 188 | | Superior only | M.A. | General Ach. | No sig. dif. | |
| Billett (1928)[a] | 9 | | 3 years | 3 | I.Q. | English | No sig. dif. Largest dif. favored AG slow pupils. | |
| Holy, Sutton (1930)[a] | 9 | | 1 sem. | | | Algebra | AG made greater gains. | |
| Fowlkes (1931) | 7 | | | | I.Q. | General Ach. | AG gained slightly less than RG on general Ach. test. | |
| Rankin (1931) | 7 | 730 | 1 year | 3 | C.A., I.Q. and Teacher Judgment | English<br>Arithmetic<br>Gen. Sci.<br>Soc. Sci. | Slow & Av. achieved more in AG. Sup. achieved more in RG. No dif. in results for dif. school subjects. | |

| Study | Grade Levels | N | Length of Study | Grouping Levels | Grouping Criteria | School Subjects Measured | Results | Remarks |
|---|---|---|---|---|---|---|---|---|
| Herr (1937) | 7-8 | 97 | 1 year | Superior Only | Past ach. I.Q. and Teacher Judgment | English History Plane Geometry, Chemistry & general information. | Dif. in American History & general information favored AG. | Matched pairs design. AG accelerated to cover 2 years work in 1. |
| Justman (1954) | 7-9 | 79 to 83 Matched Pairs | 2 years | Superior Only | I.Q. Above 130 | Math. Science Soc. Stud. | AG pupils achieved sig. more at .01 level in all 3 school subjects at end of study. No evidence that groups were comparable at beginning of study. | Matched pairs design. This study is biased in favor of AG as highly motivated pupils selected for AG leaving less motivated pupils in RG. AG completed 3 years of Jr. H. S. in 2 years. |
| Wilson (1959) | 9 | 96 Pairs Superior<br>48 Pairs Average | 1 year | 2 Superior & Av. | I.Q. & Teacher Judgment | Soc. Stud. Reading Writing, Use of Sources, Problem Solving, Critical Thinking | There were not stat. sig. dif. in gains made by pupils in AG & RG classes. | Pupils were grouped in special combined English-Soc. Studies classes only. Sup. group I.Q.'s above 115; av. group 100-110; no slow group studied. |
| Passow, Goldberg, and Link (1961) | 7-9 | 112 | 3 years | Superior Only | I.Q. | Math. | Accelerated gifted pupils achieved significantly more in mathmetics than enriched gifted pupils or control group. | Used 4 match homogeneous groups of gifted pupils. Tested effectiveness of different curriculum adjustments for homogeneous gifted groups. |
| Simpson, Martinson (1961) | 8<br>11-12 | 436 (sec.) | 1 year | Superior Only | I.Q. (S.B.) Above 130 | Arith. Reading Writing Listening Science Soc. Stud. | Special hom. classes at each grade level ach. sig. more than RG controls. | This study involves a wide variety of treatments for sup. pupils in grades 1-12. |
| Drews (1962) | 9 | 432 | 1 year | 3 | I.Q. Reading Skills, Language Skills | Problem Solving, Reading, Language, Critical Thinking | No stat. sig. dif. bet. AG & RG pupils on any of the ach. measures. | Carefully designed & executed. The only previous study which compares programs that provide extensive differentiation of the curriculum for both AG & RG pupils |

[a] Taken from Ekstrom, (1959).

Abbreviations:
AG = ability grouped
stat. = statistical, statistically
sup. = superior
dif. = difference(s), different
bet. = between
RG = random grouped
sig. = significance, significant
av. = average
hom. = homogeneous
ach. = achievement

completed by Drews (1962). Although limited in some respects in that pupils were studied at only one grade level, and one subject area, this research used exceptionally good design and assembled data on a number of important variables that have not been explored in previous research. The goal of Drews' research was to determine the effects of heterogeneous and homogeneous grouping upon pupils at three ability levels in ninth grade English. The criteria for establishing the pupils' ability level were I.Q., reading skills, and language skills. Placement of pupils into the two treatments and three ability levels was carried out using stratified random sampling prior to the start of the school year. Tests and questionnaire data were collected at the beginning and end of the year, and data were also gathered in observations, three having been made of each class involved in the study.

The final sample for whom all major data were collected included a total of 432 students. Because both the homogeneous and heterogeneous groups were divided into three ability levels, and each ability level was divided by sex for the purposes of analysis, some of the sub-samples were rather small. The smallest sub-sample in the homogeneous group was the slow girls sub-group which contained 20 cases. The same sub-group in the heterogeneous group was also smallest containing only eight cases. Class size ranged from 30 to 35 students excepting for homogeneous classes for slow students which were between 15 and 20.

Eight teachers were involved in the experiment: four experienced and four inexperienced. As a rule, each teacher taught two homogeneous classes at different levels and one heterogeneous class thus helping to cancel out possible differences in teaching ability. Experienced and inexperienced teachers taught equal proportions of the homogeneous and heterogeneous classes. All teachers met with Drews for a 90 minute period every two weeks throughout most of the experimental year in an effort to develop a common viewpoint on the methods and curriculum adjustments to be used in the homogeneous and heterogeneous classes. Teachers attempted to differentiate the curriculum in both homogeneous and heterogeneous classes. Some acceleration was carried out in the superior homogeneous classes, but the main goal of the differentiated curriculum at all ability levels was enrichment. One of the outstanding aspects of Drews' study was the great effort aimed at helping teachers develop a well-differentiated curriculum for use in both homogeneous and heterogeneous classrooms. Many previous studies have made no attempt at differentiation, (Worlton, 1928; Martin, 1927; Goldberg, et. al. 1961) thus ignoring one of the major purposes of ability grouping. Other studies have devoted a great deal of effort into developing differentiated curricula for ability grouped classes while doing nothing to provide for individual differences in random group classes used for control (Simpson and Martinson, 1961) thus introducing a considerable bias in favor of the ability grouped situation.

Initial and final achievement were measured in Drews' study using the reading and language tests from the California Achievement Battery. The advanced form was used for superior pupils, the intermediate form for average pupils, and the elementary form for slow pupils. In addition to the reading and language tests given at all ability levels, superior students were also administered tests of critical thinking and dogmatism and average pupils were administered tests of problem solving and dogmatism. The slow group was given an individual reading test in addition to the group tests of reading and language. Initial and final test scores were analyzed using the t-test to compare homogeneous and heterogeneous sub-samples of the same sex and ability levels. In reading comprehension, none of the comparable groups were significantly different on either the initial or final testing. The same comparisons of the pre- and post-tests of language achievement yielded the same results, although some of the differences obtained approached statistical significance at the .05 level. Negative results were also obtained on the critical thinking test, problem solving test, and dogmatism scale. Thus, it may be concluded from Drews' study that homogeneous grouping in ninth grade English has no significant effect upon the achievement areas measured at any of the three ability levels. In addition to the achievement data given above, Drews' study also dealt with a number of other important variable which will be discussed in the appropriate chapters of this report.

Drews' study, although very well designed, has two limitations worthy of mention. One of these, brought out by Drews herself was that the achievement measures used were aimed primarily at measurement of facts, while the type of program that Drews was attempting to stimulate might be expected to develop many aspects of achievement not measured by a fact-oriented test. The study collected other data that were aimed at measurement of these higher goals, but this weakness could have been further overcome by using an achievement battery such as the Sequential Tests of Educational Progress, which places more emphasis upon higher cognitive processes than does the California Achievement Test.

Another minor weakness of Drews' study was her use of the t-test for measuring both pre- and post-test differences, rather than the use of analysis of covariance which would have adjusted the post-test mean scores for pre-test differences that she found. Although it is doubtful that the use of analysis of covariance would have made any major change in Drews' results, it would appear from studying her data that two or three of the post-test differences would have reached statistically significance if this analysis technique had been employed. Even considering these limitations, Drews' study is exceptionally well done and probably provides more sound scientific evidence on the effects of ability grouping at the secondary level than all previous research combined.

To date a project reported by Simpson and Martinson (1961) is the most extensive study published that is concerned with special programs for superior pupils. The study was carried out by the California State Department of Education under the direction of Martinson. In this study a total of 17 different types of programs were compared, including a variety of special grouping programs, acceleration programs, and enrichment programs. Different aspects of this study were concerned with different grade levels, with all grade levels ranging from 1 to 12 being studied to some extent. Pupils who participated in the study were selected using the Stanford-Binet Scale and the study was limited to those receiving an I.Q. of 130 or higher on this measure. A total of 929 pupils were selected, about half of whom were at the elementary and half at the secondary level. A control group was set up at each grade level. This study actually employed two matching arrangements. In some situations groups were matched in such a way that group means and standard deviations were nearly identical on the matching variables which were chronological age, I.Q., sex, and socio-economic status. In other situations, matched pairs of pupils were used. The group matching techniques appear to have been more widely employed, as the numbers of cases in the experimental groups were generally larger than the numbers in the control groups at a given grade level.[1] No special program provisions were made for the control group subjects. Academic achievement was measured at the primary level with the

---

[1] The matched pairs technique produces groups of equal size.

Gates Reading Test, and two other achievement measures designed specifically for the study. Six of the Sequential Tests of Educational Progress (STEP) were administered from the fifth grade through high school, and high school seniors were also given the Graduate Record Examination Area Tests.

Overall comparisons between pupils in the experimental and control programs indicated that the first grade pupils who participated in the experimental programs made significantly greater gains in reading and arithmetic. At the upper elementary and secondary levels, the pupils in 10 out of 12 experimental programs made gains on the achievement battery that were significantly greater than those made by the control groups. Of the 17 programs involved in Martinson's research, several are directly pertinent to the area investigated in this report. Enrichment in the regular classroom is one such program. This type of program was established at first grade, fifth grade, and sixth grade levels and is comparable to the District R enrichment program described in this research report. The special class program carried out with gifted fifth and sixth grade pupils is essentially an ability grouping program similar to that employed in District A of the Utah study although superior pupils in the California program were somewhat more homogeneous and at a higher level than superior pupils as defined in the District A ability grouping system. Special ability groups for superior junior and senior high school students were established in a number of subjects. At the junior high school level these included English, social studies, and mathematics. At the high school these special superior groups were provided in English, history, algebra, and trigonometry.

The academic achievement data for pupils participating in the above programs yielded some interesting results pertinent to the Utah study. Comparisons between pupils in the fifth and sixth grade enrichment program with the control groups showed a final mean STEP score difference favoring the enrichment groups. This difference was significant beyond the .01 level. A total of 76 pupils were involved in the fifth and sixth grade enrichment program with 58 pupils in the control groups.

Differences between the control groups and the fifth and sixth grade special class program (which was essentially an ability grouped situation for superior pupils) were also statistically significant beyond the .01 level favoring the special classes over the control groups. In this case the difference in mean STEP score was higher, yielding an F-test of 22.42 which was significant beyond the .01 level. This special class sample consisted of 33 pupils. Although precise comparisons between the special class programs and the enrichment groups for fifth and sixth grades cannot be made, the data would suggest that the special class group made significantly greater gains than the enrichment group.

The eighth grade special class program involved 66 pupils. This group also achieved significantly more than the eighth grade control group with a difference between the STEP means that was significant at the .05 level. Similar results were obtained in comparisons between the eleventh grade special class program and the control groups. The 68 pupils in this special class program obtained a final STEP mean 7.4 points above the control group, this difference being significant at the .01 level.

The twelfth grade honors group program was also essentially a homogeneous ability grouped situation. This group consisted of 38 pupils. Their final STEP mean score was significantly higher than the control group, at the .05 level.

Although generally well designed, Martinson's study has several limitations. The matching techniques employed in this study constitute one of its major weaknesses. If random assignment to experimental and control groups had been possible, and admittedly this would have been extremely difficult with the large numbers of programs involved, comparisons could have been made, not only between experimental and control groups, but between the different experimental groups as well. Also, all of the usual problems and limitations attending matched group designs would have been avoided.

In some cases it appears from the available data that all pupils participating in some experimental program were to be found in a single classroom. If this were the case, a strong possibility is introduced that teacher ability and motivation could have played an important role in the results. This bias would be further aggravated if the researchers were successful in carrying out their planned design, which required that pupils selected for the control groups not be identified to their teachers during the experimental year. Thus, it appears that we have the situation found all too often in educational research in which pupils studying under teachers who are highly motivated by a new and interesting experimental program are compared with pupils studying under teachers carrying on "the same old program" and in this case not even aware of the fact that their pupils were participating as controls in a research project.

In spite of these and other limitations which will not be discussed, Martinson's research provides much valuable information and merits the careful study of anyone interested in special programs for superior students.

Another study concerned with achievement of gifted pupils has been reported by Passow, Goldberg, and Link (1961). In this study, four grouping treatments were employed for gifted pupils in mathematics. Data were collected from the seventh through the ninth grades. Twenty-eight gifted pupils were assigned to each treatment. These pupils were carefully matched on I.Q., Arithmetic Achievement, age, a teacher rating, and sex. Two of the treatments employed acceleration: one using the conventional mathematics sequence and materials while the other used materials developed by the University of Illinois Committee on School Mathematics. The enrichment class followed the regular mathematics curriculum and incorporated special enrichment units into this curriculum. The control class followed the standard mathematics curriculum. It should be noted that all of these classes were homogeneously grouped so that the comparisons made were between homogeneous grouping with acceleration, homogeneous grouping with enrichment, and homogeneous grouping with little or no provision for individual differences.

At the end of the ninth grade the STEP Mathematics Test, form 1-A, and a teacher-made mathematics test were administered to all four groups. On the STEP Mathematics Test, the accelerated group that followed the conventional program made a significantly higher mean score than did the enrichment group or the control group. Pupils in the accelerated group that were given the Illinois program were significantly higher than the control group but not significantly different from the other accelerated group or the enrichment group.

The mean scores of pupils in the four treatments on the teacher-made mathematics test were closely comparable to the STEP results with the accelerated group following the conventional program earning the highest mean score. This group was significantly higher than either the enrichment or the control group. The accelerated group that followed the Illinois program obtained the next highest mean score which was significantly higher than the control group but not different from the other accelerated group or the enrichment group. The enrichment group and the control group were not significantly different on either test. These data would lead to the conclusion that for homogeneously grouped gifted classes, acceleration is a more effective teaching technique than enrichment. Perhaps the main limitation of this study is the fact that only one class of 28 pupils was involved in each of the four treatments. As these classes were taught by different teachers, it is not possible to determine what portion of the pupil gains could be attributed to teaching ability, and what portion to the treatment differences.

## Critical Evaluation

A critical evaluation of studies related to ability grouping reveals a number of weaknesses. Perhaps the most serious of these is that many studies are of only one year duration. In most of these studies the ability grouping method is the "new" or "experimental" treatment. This situation tends to lead to a significant Hawthorne Effect which may well explain all or most of the differences found in these studies.

Another weakness of many of these studies is the use of inadequate analysis techniques. Many of the early studies and some of the more recent ones report no tests of statistical significance but merely indicate that one or the other group did "better" or "achieved more," (Moyer, 1924; Worlton, 1928; Kvaraceus and Wiles, 1938 Bonar, 1929; Roberts, 1947).

Billett (1932), in one of the early reviews of research in this field indicated that, of the studies reported between 1910 and 1928, only four made reasonable attempts to control pertinent variables. Other investigators have been content to make a single comparison between ability grouped and random grouped samples and have made no attempt to appraise possible differential effects of ability versus random grouping for pupils at different ability levels (Daniels, 1961; Woody, 1924). As there is no reason to assume that ability grouping affects superior, average, and slow pupils in the same way, these overall comparisons can have little real meaning. It is at least hypothetically possible, for example, for superior pupils to make much greater gains in ability grouped classes, average pupils to do equally well in either grouping situation, and slow pupils to make much greater gains in random grouped classes. In this case, if pupils of different ability were lumped together for analysis, the lower relative gains of the slow pupils would cancel out the higher gains of the bright pupils and lead the investigator to conclude that the two grouping systems brought about no difference in achievement. Thus, such gross comparisons can give us little insight into what really happened in an experiment.

A design weakness in many ability grouping studies is the use of matched pairs of subjects (Purdom, 1924; Jones and McCall, 1926; Martin, 1927). Many of the deficiencies of matching techniques such as the tendency for matched groups to be biased with respect to variables other than the matching variable, and the biases introduced by loss of cases because of non-random factors, were generally unknown when many of the earlier studies were carried out. These deficiencies, however, have been generally recognized by researchers for the past 15 years, but, nevertheless, a study reported in 1960 (Provust) employed matched pairs which resulted in the loss of about one-half of the cases involved in the experiment.

A number of other studies (Jones and McCall, 1926; Herr, 1937) have used techniques for selecting pupils for ability grouped classes which bias the sample and seriously distort the findings. In a study by Justman (1954), for example, pairs of intellectually gifted pupils from "special progress" and "normal progress" classes were matched on a number of variables including grade level, sex, chronological age, and mental age. Pupils for the "special progress" classes which, in this case, were homogeneous groups of superior pupils, however had been especially selected as those "...who show superior academic achievement and who also possess personal characteristics of initiative, enthusiasm, willingness to work, reliability, regular attendance, and capacity for sustained work." (Justman, 1954, p. 142).

As all of these variables have an obvious relationship to achievement, and as the pupils assigned to homogeneous "special progress" classes were especially selected for possessing these characteristics, it is hardly a surprise to learn that these pupils made significantly higher achievement scores at the end of the study than did their equally intelligent but less motivated contemporaries, who remained in the normal progress classes. These results become more suspect when we learn that no attempt was made to equate the groups on achievement at the start of the study, at which time the "special class" group was probably already superior in the achievement areas measured. Thus, although the study was well executed except for the aforementioned design errors, it seems unwise to place any confidence in Justman's study in view of these errors.

Westmeyer (1957) appears to have introduced a similar bias into his research. He compared superior pupils who elected to take a special advanced problem class in science with pupils of similarly high ability in heterogeneous science classes. Although pupils were matched on the basis of grade point average for previous high school work, it would appear that the factors that lead some pupils to elect an advanced problems course are related to motivation and interest in science and therefore would be expected to be related to science achievement. The results of this study showed that the homogeneous group selecting the advanced problems class obtained significantly higher scores at the end of the study on a standardized achievement test. It is impossible to determine, however, how much of this difference was due to ability grouping, and how much was due to sampling bias.

In evaluating research projects in this area, it is probably safe to assume that when pupils who elect to take a special class are compared with equally intelligent pupils who have not elected to take the class; or when pupils are placed in special homogeneous groups because they possess special characteristics not possessed by pupils in the control group, the sampling bias so introduced will account for much of the difference found.

While the aforementioned studies suffer principally because of the lack of sufficient experimental controls, the findings of some studies found in the ability grouping literature suffer from having been controlled too much. Ability grouping, as has been stated earlier, is essentially aimed at providing a classroom situation in which more adequate curriculum adjustments for individual differences may be made. Thus, studies of ability grouping which do not make some provisions for differentiating content and method for the different levels of ability cannot produce evidence on some of the major questions in this area. Several of the earlier studies including those carried out by Ullrich (1926); Burtt, Chassel, and Hatch (1923); and Worlton (1926) attempted to keep course content and method identical for all ability levels and for both homogeneous and heterogeneous classes in their studies. These efforts at control, of course, tend to hold constant the very area that should be manipulated in the ability grouping experiment. Such studies are valuable in answering questions about ability grouping per se, but tell us nothing about the effects of ability grouping under the conditions that it is used in the field--that is, with curriculum differentiation. Some recent studies have attempted to study the merits of ability grouping per se by developing or identifying classes which vary in terms of homogeneity and ability level and then compare achievement for these classes. A study of French (1960) employed the former technique of locating and comparing classes already in existence which happen to differ in degree of homogeneity, while the research reported by Edmiston and Benfer (1949) manipulated the degree of homogeneity while making no other adjustments. The extensive research of Goldberg et. al. (1961) discussed earlier in this chapter was also aimed at evaluating ability grouping per se.

## PROCEDURE

### Sample

Although most of the previous research on ability grouping has been concerned with achievement, the weaknesses found in many of the previous studies and their limited scope seemed to justify an extensive comparison of ability grouping and random grouping in terms of changes in achievement. Therefore, with one exception, subjects in the Utah study were administered an achievement test battery near the beginning and end of the pilot study year, and near the end of each of the subsequent three years of the study. The only exception to this testing program was that pupils in junior or senior high school during the 1960-1961 school year were not administered an achievement battery.

Achievement data were available for Samples IV, VI, VII, VIII, and IX and for a special sample of superior junior high school pupils who were the subjects of a special study. The numbers of cases involved in each specific analysis are given in the results section of this chapter. The reader will note that achievement gains from year to year as well as gains over the entire four years were studied for each sample. This procedure provided ample data to determine the nature of long term achievement performance in the two grouping treatments. Also, as achievement data for two or more samples were available at most grade levels, extensive cross checking of the results was possible.

### Achievement Measures

Because of financial limitations imposed during the pilot study year, the California Achievement Test Battery already in use in the two school districts was employed. This battery, however, was not used beyond the pilot study year for a number of reasons among which the following are perhaps the most important: (1) A test battery in regular use in the school district is more likely to be known by the teachers and more susceptible to "teaching the test" or preparing the student for specific test items. (2) The California Achievement Test Battery measures achievement by sampling specific facts and is therefore likely to be biased to some degree as a result of curriculum differences between the school districts. Although the curricula of the two school districts involved in this research were very similar--both districts selecting text books from the Utah Approved List and both districts following Utah courses of study--some specific differences certainly existed. These specific differences in the facts presented could result in the California Achievement Test Battery being better suited to the curriculum of one school district than the other; therefore favoring the former district in terms of scores. (3) The California Achievement Test Battery does not contain a sub-test in science. This was one of the two subjects in which students were ability grouped at the secondary level in District A and it was therefore necessary to employ a battery that contained a satisfactory science achievement sub-test.

A careful study of all available achievement tests was made in order to select the battery to be used during the final three years of this study. The Sequential Tests of Educational Progress (STEP) published by the Educational Testing Service, Princeton, New Jersey, were finally selected. These tests have a number of advantages which were not possessed to the same degree by any of the competing achievement batteries. Among these advantages are the following: (1) The design of the test emphasizes broad understanding and the ability to utilize learned skills in solving new problems rather than ability to handle only specific facts, thus reducing the liklihood of content validity differences for the two districts participating in the research. (2) The STEP battery contains vertically comparable tests which cover the entire grade span of subjects in this study. Many of the test batteries do not cover this wide a span and none of the alternative batteries provided as much evidence of vertical equating. Equating sub-tests which measure adjacent grade levels makes it possible to consider converted scores obtained on different levels of a given sub-test as indicating reasonably comparable levels of achievement. As most of the subjects in this research were administered two different levels of the STEP test during the four years of the study, this vertical comparability becomes a highly desirable feature. (3) Detailed studies of the norm distributions carried out by Educational Testing Service indicated that in every instance, the test provided adequate ceiling for better students in the schools from which the norm samples were drawn and in general provided an adequate floor for low ability students. This characteristic is particularly important in a study in which pupils of a given grade level but of different ability levels must be analyzed separately. The scheduling difficulties of administering two or three different levels of the achievement test at each grade level in order to obtain an adequate floor and ceiling for low and high ability pupils would have been prohibitive.

Other features of the STEP battery which are less important from a technical standpoint but of major importance from a practical standpoint include the following: (1) All STEP measures employed in this battery (mathematics, science, reading, and social studies) employ the same answer sheet and have the same time limits. The identical time limits, identical answer sheets, and similar instructions all lead to easier training of test administrators and proctors which would be expected to result in a more standard test administration situation. (2) The instructions for administering different levels and different sub-tests are very similar and in many cases identical. This greatly eases the problems of administering make-up tests when different pupils in the make-up group may simultaneously be taking different tests at different levels under the same administrator.[1]

## ACHIEVEMENT RESULTS

### Original Fourth Grade Sample (Sample IV)

#### Introduction

Most of the differences in achievement gains will be reported between pupils of comparable ability in the two districts. The analyses were designed to yield information most pertinent to the question: "Do comparable pupils in ability grouped and random grouped classrooms make comparable achievement gains?" The usual analysis procedure was to identify pupils in District A and District R who were in a given ability level classification (superior, average, or slow) and carry out a separate analysis of covariance for each pair of comparable ability groups.[1] In the case of Sample IV, however, an analysis of covariance involving all six groups (three ability levels in two districts) was also carried out. In carrying out the analyses of covariance, "t" or "F" tests were usually made between achievement gains of all combinations of groups within the given subsample, including comparisons between pupils at different ability levels within each district. These data, however, are not reported as the differential effects of regression upon the different ability groups makes interpretation difficult. In the case of the superior groups, the obtained gain in achievement includes the true gain _minus_ the effects of regression which tend to lower final achievement scores of superior groups. On the other hand, in the case of slow groups regression would be positive rather than negative so that the obtained gain equals the true gain _plus_ the regression effect. In comparing groups of similar ability level, the regression effects are still present but as they are of similar magnitude for these groups and operate in the same direction they may be expected to have little effect upon comparisons of relative performance between such groups.

---

[1] See Sequential Tests of Educational Progress, Technical Report, Catalogue Number 031-00-9, Princeton, New Jersey: Educational Testing Service, 1957, for further information.

## _First Year Achievement_

The first achievement battery was administered to Sample IV in October, 1958, when these pupils were beginning the fourth grade. This sample was studied from the beginning of the fourth grade to the end of the seventh grade and provides the best evidence that this project has to offer concerning the effects of ability grouping on the achievement of the elementary school child. Subjects were administered the _California Achievement Test, Complete Battery, Form AA_ (CAT) in October, 1958, and the _California Achievement Test, Complete Battery, Form W_, in April of 1959. The initial achievement scores earned on the _CAT Form AA_ formed the basis for establishing ability level cut off points for identifying pupils in District R whose initial achievement indicated that they were comparable to pupils assigned to one of the three ability levels in District A. An analysis of variance for the initial overall test battery scores yielded an F of 1.04 between the superior pupils in Districts A and R, .50 between average pupils and .48 between slow pupils indicating that District R pupils had been successfully classified into the three ability levels. The year's final achievement data were analyzed using analysis of covariance. Therefore, the initial achievement differences, though slight, were adjusted for in calculating the $F_{y.x}$ values shown in Table 4-3. It may be seen in Table 4-3 that comparisons of the final adjusted overall achievement means on the April test battery yielded F ratios indicating significantly higher overall performance for superior pupils in District A as compared to District R. This difference was also found in comparing average pupils. The difference in final mean achievement for slow pupils in the two districts, however, was not significant and, in this case, the trend favored the random grouped slow pupils.

---

[1] In some cases other analysis procedures were used and these are described in the appropriate section of the results.

Table 4-3

Achievement of Sample IV Pupils During Their Fourth Grade Year

| Group | Arithmetic |  |  |  | Language |  |  |  | Reading |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 80 | 4.85[a] | 6.20 | 13.55 | 80 | 5.59 | 7.04 | 20.17 | 81 | 4.93 | 6.94 | 50.35 |
| Random Grouped Superior (Dist. R) | 155 | 4.88 | 5.91 |  | 160 | 5.54 | 6.53 |  | 160 | 4.99 | 6.29 |  |
| Ability Grouped Average (Dist. A) | 86 | 4.30 | 5.39 | N.S. | 83 | 4.86 | 5.87 | 7.16 | 86 | 3.72 | 5.67 | 19.74 |
| Random Grouped Average (Dist. R) | 83 | 4.35 | 5.30 |  | 83 | 4.51 | 5.26 |  | 81 | 3.88 | 5.21 |  |
| Ability Grouped Slow (Dist. A) | 62 | 3.80 | 4.68 | N.S. | 59 | 3.70 | 4.16 | N.S. | 60 | 3.18 | 4.27 | N.S. |
| Random Grouped Slow (Dist. R) | 46 | 3.90 | 4.76 |  | 46 | 3.81 | 4.22 |  | 46 | 3.06 | 4.38 |  |

[a] CAT Grade Placement Scores

[b] All F values given are significant at the .01 level. In following tables, F or t values significant at the .05 level will be indicated *. Those not so marked are significant at the .01 level.

## Achievement in Specific Areas

In addition to comparisons of overall achievement based on the total California Achievement Test Battery, pupils were also compared on the basis of the achievement in the three sub-tests in the battery. Superior pupils in District A were significantly higher than comparable pupils in District R on all three of the sub-test scores (arithmetic, language and reading). These differences, based upon F tests between the adjusted final means, were significant at the .01 level in all three subject areas. The greatest differences were found in reading.

Comparisons of the average pupils in the two districts showed no significant differences in arithmetic. District A average pupils, however, where significantly higher at the .01 level in both the language and reading sub-tests.

Comparisons between the final adjusted mean achievement scores of slow pupils in the two districts in the three sub-tests of the CAT battery revealed no significant differences.

## A Note of Caution

In evaluating the results of the first year's achievement gains between District A and District R, several points should be kept in mind. One important factor that probably influenced the scores in District A is the fact that these data were collected during the first year that ability grouping was employed in this district. Interviews with teachers and supervisors during this year indicated that teachers had not adjusted their techniques and materials to the different ability levels as well during this year as in later years. Another factor that almost surely operated to some degree in District A during the first year is Hawthorne Effect. The ability grouping system received considerable publicity in District A and the local press favored this innovation. The change aroused interest and enthusiasm on the part of many teachers and administrators in District A. This interest which is often stimulated by introducing anything new in the schools usually can be expected to result in some favorable change in pupil performance regardless of the merits of the variable introduced.

Another factor to be considered in interpreting the first year's achievement results is that, because of the limitation in funds available during the pilot study year, the achievement data are based on the CAT which is the measure regularly used by the two districts in their own testing programs. During the pilot study year, however, the two districts employed different forms of this measure in their regular testing program. District R, which routinely administered achievement tests in the fall, used form AA of the CAT. Therefore, in order to cut the research costs, both districts were administered form AA in the fall of 1958. District A, on the other hand, routinely administered its achievement battery in the spring and had recently adopted the 1957 revision of the CAT (Forms WXYZ). Therefore, in the spring testing during the pilot study year both districts were administered Form W of the 1957 revision. If any teachers in the two districts made an effort to "teach the test" this testing sequence would have tended to raise the scores of students in District R on the initial testing as this testing used District R's form of the CAT, while in District A any such increase would have been found on the final battery scores. These two factors, operating together, would have tended to reduce the apparent achievement gains of District R pupils, and increase the apparent gains of District A pupils. Neither of these batteries was administered by district personnel and no direct evidence is available that teachers actually did prime students for the achievement test. It is reasonably safe to assume, however, that when one considers the large number of teachers involved in this research, at least a few probably attempted to "teach the test." This is not an uncommon pattern of behavior among teachers who are insecure and who often consider the achievement test results as having direct implications concerning their teaching ability.

## Second Year Achievement

During the second year of the study, pupils in Sample IV were in the fifth grade. Achievement gains for this year were computed using analysis of variance by comparing the scores on the CAT battery administered near the end of the fourth grade with a battery made up of four tests from the Sequential Tests of Educational Progress, (STEP) obtained near the end of the fifth grade. Because of the great difference in form between the CAT grade placement score and the STEP converted score, standard scores with a mean of 50 and a standard deviation of ten[1] were calculated for each of the two achievement measures. Each pupil's standard score on the initial battery was then subtracted from his standard score on the final battery in order to determine his relative gain or loss. These relative gain scores were analyzed using analysis of variance to appraise the performance of pupils of different ability levels in the two districts. The STEP and CAT have two sub-test areas in common, mathematics and reading. Gains in these two areas and on the overall battery scores were compared for this sample. The results of this analysis may be found in Table 4-4. Superior pupils did not differ significantly in mathematics gains or in overall gains based upon the composite battery scores. The superior pupils in District R, however, made significantly greater gains in reading than comparable pupils in District A. Comparisons of the average ability subsamples revealed that District R pupils made significantly greater gains in all three achievement measures (i.e., reading, mathematics, and overall battery scores). Slow pupils in District R registered significantly greater gains in reading and in the overall battery score. The difference in mathematics gains, though not statistically significant, also favored the District R sample.

These results, contradicting as they do the initial year's results, merit further study. During the first year of the research, achievement results generally favored District A. It is possible that both the superior showing of District A in the first year's analysis and the superior showing of District R in the second year are attributable to the same factor. It was mentioned earlier that the CAT, Form W administered near the end of the fourth grade was the regular District A achievement test and could have been biased in favor of District A due to "teaching the test." If this was the case, District A pupils would have had an advantage during the initial year but a disadvantage during the second year as their spuriously high achievement on the spring, 1959, battery would reduce their apparent achievement during the second year. It is not until the third year of the study that the possible biases arising from use of the district's regular achievement measures could be fully eliminated.

---

[1] Based on subjects in both districts combined.

Table 4-4

Achievement of Sample IV Pupils During Their Fifth Grade Year

| | Arithmetic | | | | Reading | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | N | Initial[a] Mean | Final[b] Mean | t | N | Initial Mean | Final Mean | t |
| Ability Grouped Superior (Dist. A) | 92 | 6.24 | 258.1 | N.S. | 91 | 6.89 | 273.7 | |
| Random Grouped Superior (Dist. R) | 100 | 6.08 | 258.0 | | 100 | 6.64 | 274.7 | 2.29* |
| Ability Grouped Average (Dist. A) | 116 | 5.43 | 246.1 | | 117 | 5.61 | 255.6 | |
| Random Grouped Average (Dist. R) | 121 | 5.46 | 250.7 | 2.75 | 121 | 5.53 | 261.8 | 4.83 |
| Ability Grouped Slow (Dist. A) | 50 | 4.52 | 237.4 | N.S. | 50 | 4.03 | 238.3 | |
| Random Grouped Slow (Dist. R) | 62 | 4.70 | 241.5 | | 63 | 4.22 | 249.9 | 3.65 |

[a] *CAT* Grade Placement

[b] *STEP* Converted Score

* Significant at .05 level, those without asterisk are significant at .01 level.

|  | Total Battery |  |  |  | Total Battery |  |  |
|---|---|---|---|---|---|---|---|
| N | Initial Mean | Final Mean | t | N | Initial Mean | Final Mean | $F_{y.x}$ |
| 85 | 6.20 | 1064.1 | N.S. | 106 | 5.06 | 6.70 | 32.98 |
| 100 | 6.58 | 1063.7 |  | 178 | 5.12 | 6.25 |  |
| 102 | 4.80 | 1010.2 |  | 103 | 4.23 | 5.59 | 7.60 |
| 121 | 5.53 | 1019.5 | 4.49 | 100 | 4.26 | 5.31 |  |
| 43 | 3.79 | 956.2 |  | 79 | 3.52 | 4.40 |  |
| 62 | 4.29 | 990.1 | 4.44 | 60 | 3.60 | 4.60 |  |
|  |  |  |  |  |  |  | N.S. |

## Third Year Achievement

During the third year of the study, pupils in this sample were in the sixth grade. For the first time, both the initial and final achievement means were based on the STEP battery. For the initial score, the four sub-tests from the STEP battery, form 4B, administered near the end of the fifth grade were used and for the final achievement score the same sub-tests from the STEP battery, form 4A, administered in April of 1961 near the end of the sixth grade were employed. It was possible during this year to analyze achievement gains on the overall battery score and also on the four sub-tests scores that were used: mathematics, science, reading, and social studies. In Table 4-5, the results of the analysis may be seen. Results for this year are based upon analysis of variance of gains between the two STEP batteries. The gains made by pupils of similar abilities in District A as compared with District R revealed no significant differences in mathematics, reading, social studies, or in the overall STEP battery score obtained by combining the four sub-test scores. Only in science did statistically significant differences occur. As may be seen in Table 4-5, these differences were present at all three ability levels and in every case favored District A. An examination of the mean science gains indicates that the greatest gains were made by pupils in the slow subsample with next greater gains occurring in the average group and smallest gains among pupils in the superior group. This trend is probably due to regression which works in favor of the slow group and against the superior group. A less likely explanation would be that the ceiling of the STEP test may have been too low to provide an adequate measure of the superior sixth grade pupils. In the four sub-tests, the mean score for the superior groups on the STEP form 4A taken near the end of the sixth grade ranged between the 79th percentile in reading and the 89th percentile in mathematics.[1] The normative data were obtained in the fall of the sixth grade while our subjects were tested in the spring of the sixth grade. If spring norms were available, these percentiles would be lower. In any case, however, the superior group does not appear to be seriously restricted by the ceiling of the STEP tests. This ceiling would, of course, be less restrictive for these pupils when they were in the fourth and fifth grades.

## Fourth Year Achievement

During the fourth year of the study, pupils in Sample IV were in the seventh grade. It will be recalled that at the seventh grade level, District A employed ability grouping only in mathematics and science. Table 4-6 gives the results of analysis of variance of gains in each of the four areas of the STEP battery. The STEP battery, form 4A, administered near the end of the sixth grade was used as the initial measure and the STEP battery form 3B, administered near the end of the seventh grade, as the final measure. Superior pupils in both districts were given the high school form of the STEP (2B) in order that there would be no chance of their being restricted by the test ceiling. Analysis of variance was not carried out to appraise overall achievement gains in view of the fact that junior high school students in District A were not ability grouped in reading and social studies and therefore an overall score would have no meaning in terms of the goals of the research.

---

[1] See Appendix D of the *STEP Technical Report*, Princeton, N.J.: Educational Testing Service, 1957.

The results revealed only one significant difference in comparisons between the ability grouped and random grouped seventh grade pupils of similar ability levels on the STEP mathematics and science tests. It was found that superior ability grouped pupils made significantly greater gains in mathematics than superior random grouped pupils. In spite of the fact that no regular grouping procedure was followed in District A in English and social studies, it appears likely that classes in these subjects were somewhat more homogeneous than were classes in District R. This would be expected because the process of ability grouping pupils in two school subjects brings about a scheduling situation in which some grouping is almost certain to occur in other subjects as well. It is interesting to note that, although not statistically significant, the mean gains favored District A superior pupils in all four comparisons with the District R superior pupils. Conversely, the District R slow pupils made greater achievement gains in all four sub-tests than the District A slow pupils although again these differences were not large enough to be statistically significant. For average pupils, the differences between the two districts were not consistent and were generally small.

### Overall Results for the Four Years

In addition to analyzing gains year by year for this sample, the scores on the initial CAT battery administered in the fall of 1958 were compared with the scores of the same pupils on the final STEP battery administered in spring of 1962. Scores on the two measures were converted to standard scores combining the students in both districts to establish the mean and standard deviation used in the conversion. Analysis of variance was computed on relative standard score gains for the six groups between the initial and final batteries. These results indicated that for the four years, the ability grouped superior sample had made significantly greater overall gains. For the average groups, there was no significant difference, and for the slow groups, the District R pupils made significantly greater gains. For this comparison the pupils were classified in accordance with their ability level at the outset of the study while in the year to year comparisons, classification was based upon the pupils' ability level at the beginning of the year in question. The year by year analyses appear to have one disadvantage. They fail to reflect adequately the cumulative effects of the two grouping systems as pupils shift in their ability level classification from year to year if their performance justifies a change. For example, a child who is initially in the slow group in District A may be moved to the average group for the second year of the study and will thus become part of the average group during the second year's analysis. There was, in fact, a rather large number of pupils in District A who did move out of the slow group. The data for the entire four years fill in this gap, and provide a more accurate picture of long term gains. A limitation of the four year data, however, is its inclusion of the first year achievement gains which may have been influenced by Hawthorne Effect in District A. It will be recalled, however, that the various ability groups were very similar on the initial CAT battery (fall, 1958) so that the four year gains seem to offer a sound basis for evaluating the overall achievement effects of the two systems.

### Results for the Last Three Years

If Hawthorne Effect were a major factor during the first year, an analysis of achievement for the last three years of the study might give a more accurate picture of relative

Table 4-5

Achievement of Sample IV Pupils During Their Sixth Grade Year

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean[a] | Final[b] Mean | t | N | Initial Mean | Final Mean | t |
| Ability Grouped Superior (Dist. A) | 82 | 257.85 | 263.12 | N.S. | 82 | 268.90 | 272.12 | 2.07 |
| Random Grouped Superior (Dist. R) | 220 | 258.25 | 264.39 | | 222 | 271.24 | 271.99 | |
| Ability Grouped Average (Dist. A) | 135 | 246.36 | 253.78 | N.S. | 137 | 256.99 | 263.76 | 2.44 |
| Random Grouped Average (Dist. R) | 152 | 248.99 | 254.49 | | 152 | 258.30 | 262.71 | |
| Ability Grouped Slow (Dist. A) | 42 | 237.61 | 242.97 | N.S. | 42 | 242.54 | 251.71 | 2.09 |
| Random Grouped Slow (Dist. R) | 80 | 237.90 | 245.14 | | 80 | 246.70 | 251.95 | |

[a] STEP Converted Score

[b] In calculating overall achievement, the total score for pupils lacking one of the four STEP scores was multiplied by 4/3 and included in the analysis. These cases were omitted in the analysis of achievement in the specific areas.

| Reading | | | | Social Studies | | | | Total Battery | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N | Initial Mean | Final Mean | t | N | Initial Mean | Final Mean | t | N[a] | Initial Mean | Final Mean | t |
| 81 | 274.05 | 275.54 | N.S. | 80 | 263.72 | 266.85 | N.S. | 85 | 1064.03 | 1078.00 | N.S. |
| 231 | 276.63 | 277.78 | | 233 | 264.67 | 267.73 | | 245 | 1070.92 | 1080.77 | |
| 140 | 256.06 | 263.97 | N.S. | 141 | 252.70 | 256.48 | N.S. | 148 | 1012.41 | 1035.09 | N.S. |
| 163 | 257.35 | 263.30 | | 159 | 253.59 | 256.57 | | 175 | 1018.34 | 1037.21 | |
| 43 | 239.75 | 247.02 | N.S. | 40 | 240.32 | 246.58 | N.S. | 45 | 961.42 | 989.93 | N.S. |
| 82 | 243.99 | 248.19 | | 81 | 242.10 | 247.16 | | 84 | 970.69 | 992.68 | |

Table 4-6

Achievement of Sample IV Pupils During Their Seventh Grade Year

| | Mathematics | | | | | Science | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Group | N[a] | Initial Mean | Final Mean | Mean Gain | t | N[b] | Initial Mean | Final Mean | Mean Gain | t |
| Ability Grouped Superior (Dist. A) | 75 | 263.88 | 274.41 | 10.53 | | 67 | 272.31 | 281.07 | 8.76 | |
| Random Grouped Superior (Dist. R) | 212 | 264.44 | 272.56 | 8.12 | 2.21 | 212 | 271.88 | 278.08 | 6.20 | N.S. |
| Ability Grouped Average (Dist. A) | 145 | 256.26 | 261.54 | 6.28 | | 187 | 265.30 | 270.06 | 4.76 | |
| Random Grouped Average (Dist. R) | 172 | 254.50 | 262.25 | 7.75 | N.S. | 149 | 262.76 | 268.67 | 5.91 | N.S. |
| Ability Grouped Slow (Dist. A) | 37 | 241.57 | 246.35 | 4.78 | | 31 | 250.91 | 258.19 | 7.28 | |
| Random Grouped Slow (Dist. R) | 78 | 245.27 | 252.89 | 7.62 | N.S. | 78 | 252.45 | 261.14 | 8.69 | N.S. |

[a] Different criteria used in District A to group in math and science on secondary level, thus N's slightly different. Most pupils were in same ability level for both subjects.

[b] As neither district was ability grouped in English or Social Studies, the 6th grade grouping levels were used to classify pupils in these subjects.

|  | Reading |  |  |  |  | Social Studies |  |  |  |
|---|---|---|---|---|---|---|---|---|---|
| N$^b$ | Initial Mean | Final Mean | Mean Gain | t | N | Initial Mean | Final Mean | Mean Gain | t |
| 76 | 275.96 | 288.02 | 12.06 | N.S. |  | 266.91 | 272.27 | 5.36 | N.S. |
| 212 | 277.92 | 287.65 | 9.73 |  |  | 267.74 | 271.28 | 3.54 |  |
| 131 | 264.16 | 272.00 | 7.84 | N.S. |  | 256.70 | 259.84 | 3.14 | N.S. |
| 149 | 253.14 | 271.77 | 8.63 |  |  | 256.51 | 258.79 | 2.28 |  |
| 40 | 249.37 | 257.88 | 8.51 | N.S. |  | 247.99 | 250.76 | 2.77 | N.S. |
| 78 | 248.02 | 259.45 | 11.43 |  |  | 247.11 | 251.59 | 4.48 |  |

achievement. An analysis of variance was therefore carried out between overall scores on the CAT battery administered at the end of grade four and the STEP battery administere at the end of grade seven. This analysis revealed no significant difference between superior pupils in the two districts, suggesting that the four year differences were due primarily to the large first year gains. A significant difference favoring the District R average sample was found. The District R slow sample also made significantly greater gains over the final three years of the study.

## Conclusions

When one considers all data comparing the superior pupils in the two grouping systems it must be concluded that no evidence is present to indicate that either system is consistently more effective in promoting achievement. The evidence, however, tends to favor the ability grouping system in that of the 15 comparisons made, over the four years 12 were in favor of District A superior pupils with two of these significant at the .01 level and two significant at the .05 level. Only three comparisons favored District R pupils with one of these significant at the .05 level. The four year results show a significant difference in favor of District A superior pupils, but as this differenc is not present in the final three-year results one must conclude it is due primarily to the very large first year gains made by the District A superior pupils.

The overall picture for pupils of average ability shows no definite trend with eight difference (four significant) favoring District R and seven differences (two significant) favoring District A.

The evidence is somewhat stronger with respect to slow pupils. Eleven of the fifteen differences found favored District R including four of the five statistically significant differences. The overall results for the entire four years are significant and favor District R as do the overall results for the final three years of the study. On the basis of these results it appears justifiable to conclude that significantly greater achievement might be expected from slow pupils in random grouped elementary school classrooms as compared with pupils of comparable ability in ability grouped classrooms. These results for the slow pupil must be considered as at least partially due to initial intelligence differences. The District R slow sample was found to have a higher verbal intelligence grade placement than the District A slow sample (Table 4-6). As verbal intelligence correlates rather highly with general academic achievement, we can expect much of this difference to be compensated for in the covariance analysis.

## Original Sixth Grade Sample (Sample VI)

This sample was tested at the beginning of the sixth grade in October 1958 and was followed through the ninth grade with final achievement measures administered in April, 1962.

## First Year Achievement

Achievement during the sixth grade was measured by the California Achievement Test Battery. Form AA was administered near the start of the year, and Form W was administered near the end. The same limitations apply to these data that were discussed previously with regard to Sample IV. It may be seen on Table 4-7 that the initial mean scores of District A and District R pupils at the three ability levels were closely comparable. Analysis of variance of the initial total battery scores yielded F values of .24, .76, and .52 between slow, average and superior pupils in the two districts. The final overall achievement means were adjusted for these slight initial differences using covariance analysis. Adjusted final means show that superior pupils in District A made considerably greater gains on the CAT composite battery score than their peers in District R. Differences significant beyond the .01 level were also found between average pupils in the two districts, again favoring District A. No significant differences were found between slow pupils in the two districts. The pattern of differences for this sample is very similar to that found in the first year performance of Sample IV (see Table 4-3).

## Specific Achievement Areas

Achievement on the three specific sub-tests of the CAT: arithmetic, language and reading were also compared. These results also followed a pattern generally quite similar to that found in the Sample IV analysis. District A superior pupils had significantly higher final adjusted mean scores on all three of the CAT sub-tests. These differences were all significant beyond the .01 level with the arithmetic and reading differences amounting to an over one-half year in terms of grade placement scores and yielding very large adjusted F's (90.82 and 103.64 respectively) on the analysis of variance. It will be noted that the greatest difference appeared in reading; the same results found with Sample IV.

A comparison of the final adjusted mean scores of average pupils in the two districts revealed significant differences favoring District A pupils in arithmetic and reading. These differences were much smaller than those found in the comparisons of superior pupils. No significant differences were found between the average pupils in the two districts on final adjusted language test mean score.

The District A slow pupils were significantly higher in their final adjusted language means than comparable District R pupils. There were no significant differences between slow pupils in the two districts in arithmetic or reading.

## Second Year's Achievement

In order to determine achievement gains during the seventh grade, the CAT scores earned by Sample VI pupils at the end of the sixth grade were used to adjust the means on the STEP mathematics and science tests that were administered at the end of the seventh grade. At the secondary level, District A employed ability grouping in mathematics and science only, so comparisons between the two districts will be limited to these subjects. The CAT arithmetic score was used as the initial variable in the mathematics covariance analysis.

The results showed the District R superior and slow pupils to be significantly higher than comparable District A pupils in mathematics at the end of grade seven. This difference, though not large for the superior sample, is noteworthy as it cancels out the rather large difference favoring the District A pupils that was found during the first year of the study. It will be noted (Table 4-8) that the STEP mathematics scores earned by the two samples at the end of grade seven are less than one point apart, indicating that, when one considers the first two years for this sample, achievement in mathematics was nearly identical for superior pupils in the two districts. In contrast, the significant final mean differences between the slow pupils leave the District R

Table 4-7

Achievement of Sample VI Pupils During Their Sixth Grade Year

| | Arithmetic | | | | Language | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean[a] | Final Mean[a] | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 125 | 6.55[a] | 9.52 | 90.82 | 125 | 7.10 | 8.55 | 14.62 |
| Random Grouped Superior (Dist. R) | 194 | 6.40 | 8.48 | | 197 | 7.08 | 8.20 | |
| Ability Grouped Average (Dist. A) | 114 | 5.83 | 7.62 | 5.10[a] | 117 | 6.29 | 7.37 | N.S. |
| Random Grouped Average (Dist. R) | 119 | 5.82 | 7.33 | | 125 | 6.08 | 7.05 | |
| Ability Grouped Slow (Dist. A) | 110 | 5.11 | 6.40 | N.S. | 108 | 5.28 | 6.07 | 13.50 |
| Random Grouped Slow (Dist. R) | 69 | 5.15 | 6.28 | | 65 | 5.24 | 5.60 | |

[a] CAT Grade Placement

| Reading | | | | Total Battery | | | |
|---|---|---|---|---|---|---|---|
| N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| 125 | 7.01 | 8.77 | 103.64 | 141 | 6.77 | 8.96 | 41.03 |
| 196 | 7.21 | 8.02 | | 244 | 6.90 | 8.23 | |
| 116 | 5.81 | 7.65 | 29.95 | 136 | 5.79 | 7.52 | 20.4 |
| 124 | 5.63 | 7.04 | | 153 | 5.84 | 7.14 | |
| 109 | 4.66 | 6.07 | N.S. | 128 | 4.95 | 6.13 | N.S. |
| 65 | 4.57 | 5.88 | | 88 | 4.92 | 5.97 | |

sample somewhat ahead in mathematics achievement for the initial two years.

STEP science scores were adjusted for initial differences using the CAT composite battery score. No significant differences in the adjusted final STEP science mean scores were found between any of the comparable ability groups in the two districts.

### Third and Fourth Year's Achievement

No achievement measures were administered to the secondary school sample during the third year of the study, so the achievement during the final two years of the study is covered in one analysis. Gains made between the end of grade seven and the end of grade nine on the STEP achievement battery are compared using analysis of covariance. It will be recalled that District A employed ability grouping in mathematics and science only at the secondary level, so comparisons were confined to these achievement areas. It may be seen in Table 4-9 that no significant differences were found between comparable ability groups in mathematics. A small difference favoring the District R average sample was found in analysis of the science gains. Gains at the superior and slow ability levels were not significantly different for the two grouping systems.

### Overall Results for the Four Years

As three of the four years covered in this analysis involved secondary school achievement, analysis of gains in mathematics and science seemed more appropriate than an analysis of overall gains such as was used in the four-year analysis of Sample IV achievement. Thus, two covariance analyses were conducted using CAT arithmetic grade placement and composite CAT grade placement as the initial scores and STEP mathematics and science tests as the final scores in mathematics and science respectively. The results may be seen in Table 4-10. In mathematics, the superior District A sample showed a small but significantly greater gain over the four year period. The District A average pupils registered a similar gain in science over their District R peers.

It should be remembered, however, that the final science means were adjusted in this case using the composite CAT score as there is no science test in the CAT battery. A science pre-test would, of course, have permitted a more accurate adjustment for initial differences, but the data indicate that the adjustment used was reasonable satisfactory.

## Original Seventh Grade Sample (Sample VII)

This sample was tested at the beginning of the seventh grade in October, 1958 and was followed through the tenth grade with final achievement measures administered in April, 1962. Along with Sample VI, these pupils represent the best data provided in this research concerning the effects of ability grouping at the junior high school level. During the three of the four years of this study, pupils in Samples VI and VII were at the junior high school level.

### First Year's Achievement

The first year achievement data for this sample are based upon scores on the California Achievement Test Battery. It will be recalled that ability grouping at the secondary level in District A was carried out only in mathematics and science. The

Table 4-8

Achievement of Sample VI Pupils During Their Seventh Grade Year

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 101 | 9.49[a] | 272.66[b] | 4.01* | 69 | 9.28 | 283.16 | N.S. |
| Random Grouped Superior (Dist. R) | 110 | 9.19 | 273.45 | | 87 | 8.88 | 282.13 | |
| Ability Grouped Average (Dist. A) | 211 | 7.47 | 259.73 | N.S. | 242 | 7.43 | 270.70 | N.S. |
| Random Grouped Average (Dist. R) | 178 | 7.66 | 262.11 | | 212 | 7.61 | 273.00 | |
| Ability Grouped Slow (Dist. A) | 39 | 6.13 | 247.44 | 11.69 | 42 | 5.86 | 260.69 | N.S. |
| Random Grouped Slow (Dist. R) | 90 | 6.03 | 253.56 | | 93 | 5.96 | 263.94 | |

[a] CAT Grade Placement Scores
[b] STEP Converted Scores

## Table 4-9

Achievement of Sample VI Pupils During Their Eighth and Ninth Grade Years

| Group | Mathematics |  |  |  | Science |  |  |  |
|---|---|---|---|---|---|---|---|---|
|  | N | Initial Mean[a] | Final Mean | t | N | Initial Mean | Final Mean | t |
| Ability Grouped Superior (Dist. A) | 134 | 272.45[a] | 282.58 | N.S. | 62 | 282.94 | 290.90 | N.S. |
| Random Grouped Superior (Dist. R) | 91 | 271.85 | 283.86 |  | 70 | 281.01 | 287.77 |  |
| Ability Grouped Average (Dist. A) | 309 | 259.72 | 270.89 | N.S. | 397 | 270.72 | 279.42 | 2.26* |
| Random Grouped Average (Dist. R) | 141 | 262.59 | 273.84 |  | 175 | 272.91 | 279.41 |  |
| Ability Grouped Slow (Dist. A) | 44 | 248.64 | 256.34 | N.S. | 32 | 260.03 | 267.03 | N.S. |
| Random Grouped Slow (Dist. R) | 74 | 252.91 | 259.39 |  | 74 | 263.96 | 270.74 |  |

[a] STEP Converted Scores

Table 4-10

Achievement of Sample VI Pupils Over a Four-Year Period

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 97 | 6.51[a] | 281.87[b] | 5.82* | 64 | 6.85 | 290.55 | N.S. |
| Random Grouped Superior (Dist. R) | 93 | 6.49 | 284.94 | | 73 | 7.14 | 288.37 | |
| Ability Grouped Average (Dist. A) | 182 | 5.72 | 270.40 | N.S. | 216 | 5.86 | 279.85 | 4.46* |
| Random Grouped Average (Dist. R) | 147 | 6.00 | 273.89 | | 186 | 6.23 | 279.43 | |
| Ability Grouped Slow (Dist. A) | 29 | 4.84 | 254.69 | N.S. | 35 | 4.69 | 266.71 | N.S. |
| Random Grouped Slow (Dist. R) | 78 | 5.28 | 259.73 | | 79 | 5.22 | 271.15 | |

[a] CAT Grade Placement Scores
[b] STEP Converted Scores

Table 4-11

Achievement of Sample VII Pupils During Their Seventh Grade Year

| Group | N | Mathematics |  |  |
|---|---|---|---|---|
|  |  | Initial Mean[a] | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 36 | 8.28 | 9.86 | N.S. |
| Random Grouped Superior (Dist. R) | 37 | 7.94 | 9.78 |  |
| Ability Grouped Average (Dist. A) | 29 | 6.61 | 7.74 | 4.96* |
| Random Grouped Average (Dist. R) | 57 | 6.64 | 8.36 |  |
| Ability Grouped Slow (Dist. A) | 28 | 5.89 | 6.27 | 9.50 |
| Random Grouped Slow (Dist. R) | 32 | 5.91 | 7.08 |  |

[a] CAT Grade Placement

first year achievement analysis for this sample, however, was limited to mathematics as no science measure is included in the CAT battery. Analysis of covariance was carried out for mathematics achievement, adjusting the final CAT grade placement mean score for differences in the initial mean score. The results may be found in Table 4-11. It may be seen that comparisons of superior pupils yielded no significant differences. District R average pupils showed significantly higher final means on the CAT arithmetic test than comparable District A pupils. A statistically significant and somewhat larger difference was also found between the final achievement means of slow pupils with District R pupils again showing superior achievement.

### Second Year's Achievement

Achievement during the eighth grade for Sample VII was determined by adjusting end-of-year STEP achievement means in mathematics and science for initial differences found in the California Achievement Test Battery administered at the end of the seventh grade. As with Sample VI, the overall CAT battery score was used as the initial variable in the science covariance analysis. Comparisons between comparable groups in District A and R in final adjusted mean scores in mathematics showed a significant difference in favor of the District A superior pupils. Comparisons between average and slow pupils in mathematics achievement yielded no significant differences (see Table 4-12).

The final adjusted mean scores in science also showed statistically significant differences in favor of the District A superior pupils. District A average pupils were also significantly higher than the comparable group in District R on this final achievement mean. No significant differences were found between slow pupils in the two districts.

### Third and Fourth Year's Achievement

Achievement of Sample VII pupils between the end of grade eight and the end of grade ten probably provides the most reliable comparative data for this sample. These data include adequate numbers of cases, extend over a two-year period and compare gains on two different forms of the STEP mathematics and science tests rather than making comparisons between the STEP and CAT such as were involved in the previous year. As may be seen in Table 4-13 comparisons in mathematics yielded no significant differences between pupils of comparable ability in the two districts. In the final adjusted science achievement means the District R slow pupils were significantly higher. No significant differences were found between superior or average pupils in science achievement.

### Overall Results for the Four Years

The overall gains in mathematics and science achievement for the four years of the study may be found in Table 4-14. These data, although contributing to the total achievement picture, are limited to some extent as compared with the third and fourth year analysis because of the necessity of adjusting the final STEP means for initial differences on the CAT battery and because of the somewhat smaller number of pupils for whom data were available over the four-year period. For the most part, however, the sub-groups involved in the four-year analysis for Sample VII are of adequate size, and give us some worthwhile evidence of the cumulative effect of ability grouping vs.

Table 4-12

Achievement of Sample VII Pupils During Their Eighth Grade Year

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 17 | 10.42[a] | 278.53[b] | 6.18* | 18 | 10.25 | 288.61 | 6.39* |
| Random Grouped Superior (Dist. R) | 29 | 10.20 | 271.69 | | 33 | 9.89 | 278.70 | |
| Ability Grouped Average (Dist. A) | 36 | 8.06 | 264.92 | N.S. | 36 | 7.64 | 275.11 | 6.04* |
| Random Grouped Average (Dist. R) | 60 | 8.05 | 263.73 | | 46 | 7.95 | 271.65 | |
| Ability Grouped Slow (Dist. A) | 18 | 5.68 | 248.28 | N.S. | 17 | 6.15 | 262.82 | N.S. |
| Ability Grouped Slow (Dist. R) | 21 | 5.95 | 249.54 | | 21 | 6.03 | 264.24 | |

[a] CAT Grade Placement Scores

[b] STEP Converted Scores

Table 4-13

Achievement of Sample VII Pupils During Their Ninth and Tenth Grade Years

| Group | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 95 | 278.93[a] | 283.19 | N.S. | 96 | 286.49 | 290.93 | N.S. |
| Random Grouped Superior (Dist. R) | 122 | 278.82 | 284.46 | | 84 | 288.51 | 289.71 | |
| Ability Grouped Average (Dist. A) | 209 | 266.66 | 272.99 | N.S. | 210 | 275.34 | 279.72 | N.S. |
| Random Grouped Average (Dist. R) | 87 | 265.43 | 272.08 | | 98 | 273.72 | 281.43 | |
| Ability Grouped Slow (Dist. A) | 38 | 249.66 | 262.13 | N.S. | 30 | 260.37 | 266.33 | 7.21 |
| Random Grouped Slow (Dist. R) | 44 | 248.66 | 258.68 | | 75 | 263.21 | 272.88 | |

[a] STEP Converted Scores

## Table 4-14

### Achievement of Sample VII Pupils Over a Four-Year Period

| Group | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 29 | 8.37[a] | 277.76[b] | N.S. | 30 | 8.81 | 289.43 | N.S. |
| Random Grouped Superior (Dist. R) | 54 | 8.64 | 278.18 | | 57 | 8.95 | 285.68 | |
| Ability Grouped Average (Dist. A) | 26 | 6.47 | 271.92 | N.S. | 26 | 7.09 | 278.35 | N.S. |
| Random Grouped Average (Dist. R) | 46 | 7.01 | 274.15 | | 49 | 7.29 | 278.35 | |
| Ability Grouped Slow (Dist. A) | 23 | 5.85 | 260.83 | N.S. | 19 | 5.86 | 270.79 | N.S. |
| Random Grouped Slow (Dist. R) | 28 | 5.94 | 259.36 | | 28 | 6.02 | 268.57 | |

[a] CAT Grade Placement Scores
[b] STEP Converted Scores

random grouping on mathematics and science achievement.

No significant achievement differences were found between District A and District R at any of the three ability levels in either science or mathematics. The only cumulative four year difference that approached statistical significance was between the final adjusted science means of superior pupils. This covariance analysis yielded an F-test of 3.41, and favored the District A sample.

## Original Eighth Grade Sample (Sample VIII)

This sample was tested at the beginning of the eighth grade in October, 1958, and was followed through the eleventh grade with final achievement measures administered in April, 1962. As was the case with Sample VII, the first year's achievement analysis for Sample VIII was limited to mathematics achievement. End-of-year achievement based on the CAT arithmetic test was adjusted for initial achievement differences on a different form on the same test.

### First Year's Achievement

Comparisons of the final adjusted means indicated a significant difference in favor of District R average pupils. Differences between District A and District R superior and slow pupils were not statistically significant (see Table 4-15).

### Second Year's Achievement

During the second year of the study, Sample VIII pupils were in the ninth grade. Achievement at the end of the ninth grade was measured using the STEP mathematics and science test and the obtained means were adjusted for initial differences found on the CAT administered at the end of the eighth grade following the same procedures used in Sample VII. Results of the covariance analysis may be found in Table 4-15. It may be seen that District A superior and slow pupils obtained significantly higher adjusted final mean scores in mathematics than comparable District R pupils. No significant differences were found between average pupils in the two districts in mathematics achievement. In comparing science achievement, a significant difference was found between the average groups, with the District A average pupils making greater gains. No significant differences were found between superior or slow pupils in science achievement.

### Third and Fourth Year's Achievement

The achievement made by Sample VIII pupils between the end of the ninth grade and the end of the eleventh grade in mathematics and science was measured by two forms of the STEP achievement test in these areas. Because of the wide range of electives available during the high school years, these data can be considered only tentative evidence on the relative merits of the two grouping systems. The minimum requirements in both mathematics and science were the same in the two districts, but as students could meet the minimum requirements any time during high school and with any of a variety of courses, it was not possible to obtain enough cases with identical mathematics or science programs to make precise comparisons. It may be seen in

Table 4-15

Achievement of Sample VIII Pupils During Their Eighth and Ninth Grade Years

| | Eighth Grade | | | | Ninth Grade | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mathematics | | | | Mathematics | | | | Science | | | |
| Group | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean[a] | Final Mean[b] | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 64 | 9.76[a] | 10.84 | N.S. | 63 | 10.79[a] | 281.46 | 4.35* | 55 | 10.69 | 286.47 | N.S. |
| Random Grouped Superior (Dist. R) | 76 | 9.26 | 10.72 | | 68 | 10.70 | 278.68 | | 70 | 10.65 | 284.56 | |
| Ability Grouped Average (Dist. A) | 49 | 7.65 | 8.45 | | 62 | 7.97 | 263.44 | N.S. | 60 | 8.63 | 280.62 | |
| Random Grouped Average (Dist. R) | 66 | 7.58 | 8.93 | 4.62* | 45 | 8.80 | 264.89 | | 45 | 8.72 | 274.58 | 10.34 |
| Ability Grouped Slow (Dist. A) | 54 | 6.04 | 6.73 | N.S. | 24 | 6.91 | 254.42 | 5.52* | 33 | 6.26 | 259.91 | N.S. |
| Random Grouped Slow (Dist. R) | 51 | 6.06 | 6.99 | | 32 | 7.13 | 248.34 | | 33 | 6.89 | 260.45 | |

[a] CAT Grade Placement Score
[b] STEP Converted Score

Table 4-16 that no significant differences were found between the final adjusted means in either mathematics or science between District A and District R pupils at any of the three ability levels.

### Overall Results for the Four Years

A comparison of achievement of Sample VIII pupils over the entire four years of the study may be found in Table 4-17. It may be seen that none of the differences in final adjusted means in mathematics or science between comparable District A and District R pupils were statistically significant. Two of the differences obtained approached significance at the .05 level. District A slow pupils made somewhat greater gains in mathematics achievement ($F_{y.x} = 3.67$), and the greater gains of District A average pupils in science nearly reached the .05 level ($F_{y.x} = 3.95$). However, when it is remembered that these differences represent the acculative effects of four years under the ability grouping and random grouping systems, the lack of any differences reaching the .05 level certainly suggests that the two systems had little or no differential effect upon the long term achievement gains of these pupils.

## Original Ninth Grade Sample (Sample IX)

This sample was tested at the beginning of the ninth grade in October, 1958, and was followed through the end of the twelfth grade when final achievement measures were administered. This group is somewhat smaller than samples at the other grades but represents the best evidence available in this research concerning the relative achievement gains of ability grouped and random grouped pupils in mathematics and science at the high school level. During three of the four years of the study, these pupils were high school students. The same limitations that are imposed by the range of mathematics and science electives at the high school level also apply to the Sample IX achievement results.

### First Year's Achievement

Achievement data during the ninth grade was limited to mathematics achievement on the *California Achievement Test*. Final *CAT* mathematics mean scores on form W were adjusted for initial differences on the *CAT* form AA, using covariance analysis. It may be seen in Table 4-18 that no significant differences were found between the ability grouped and randomed grouped pupils on mathematics achievement during this year. It will be recalled that a larger percentage of District A pupils take mathematics in the ninth grade, although requirements for the two districts do not differ.

### Second Year's Achievement

Analysis for the second year of the study was based on the *CAT* battery administered at the end of the ninth grade and the *STEP* mathematics and science tests administered at the end of the tenth grade. As has been the case with previous samples discussed, the final *STEP* mathematics mean scores were adjusted for initial differences found on the *CAT*. Results of the second year's analysis may be found in Table 4-18. It may be seen that no significant differences were found between District A and District R

Table 4-16

Achievement of Sample VIII Pupils During Their Tenth and Eleventh Grade Years

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 97 | 281.01[a] | 286.39 | N.S. | 83 | 289.61 | 294.07 | N.S. |
| Random Grouped Superior (Dist. R) | 114 | 281.39 | 287.05 | | 95 | 293.51 | 292.27 | |
| Ability Grouped Average (Dist. A) | 160 | 268.17 | 274.71 | N.S. | 167 | 277.71 | 282.57 | N.S. |
| Random Grouped Average (Dist. R) | 50 | 269.92 | 272.68 | | 37 | 276.84 | 282.38 | |
| Ability Grouped Slow (Dist. A) | 37 | 258.00 | 268.54 | N.S. | 35 | 262.48 | 274.06 | N.S. |
| Random Grouped Slow (Dist. R) | 41 | 250.54 | 261.37 | | 70 | 265.01 | 275.74 | |

[a] STEP Converted Score

Table 4-17

Achievement of Sample VIII Pupils Over a Four-Year Period

| Group | Mathematics |  |  |  | Science |  |  |  |
|---|---|---|---|---|---|---|---|---|
|  | N | Initial Mean[a] | Final Mean[b] | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 51 | 9.92 | 286.92 | N.S. | 53 | 9.63 | 291.68 | N.S. |
| Random Grouped Superior (Dist. R) | 60 | 9.14 | 284.50 |  | 59 | 9.49 | 288.58 |  |
| Ability Grouped Average (Dist. A) | 32 | 7.85 | 277.47 | N.S. | 32 | 7.80 | 286.56 | N.S. |
| Random Grouped Average (Dist. R) | 49 | 7.50 | 272.20 |  | 49 | 7.86 | 280.41 |  |
| Ability Grouped Slow (Dist. A) | 32 | 6.12 | 269.47 | N.S. | 30 | 6.08 | 275.87 | N.S. |
| Random Grouped Slow (Dist. R) | 19 | 6.15 | 261.57 |  | 20 | 6.19 | 271.45 |  |

[a] CAT Grade Placement Score

[b] STEP Converted Score

Table 4-18

Achievement of Sample IX Pupils During Their Ninth and Tenth Grade Years

| | Ninth Grade | | | | Tenth Grade | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mathematics | | | | | Mathematics | | | | Science | | | |
| Group | N | Initial Mean[a] | Final Mean | $F_{y.x}$ | N | Initial Mean[a] | Final Mean[b] | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 32 | 10.26[a] | 11.20 | N.S. | 29 | 11.03 | 281.72 | N.S. | 25 | 11.03 | 289.68 | N.S. |
| Random Grouped Superior (Dist. R) | 33 | 10.46 | 11.56 | | 49 | 11.61 | 285.35 | | 51 | 11.26 | 288.14 | |
| Ability Grouped Average (Dist. A) | 32 | 8.42 | 9.48 | N.S. | 72 | 9.11 | 271.28 | N.S. | 68 | 9.35 | 282.19 | N.S. |
| Random Grouped Average (Dist. R) | 33 | 8.80 | 9.92 | | 55 | 9.74 | 273.15 | | 54 | 9.44 | 280.89 | |
| Ability Grouped Slow (Dist. A) | 29 | 7.23 | 8.23 | N.S. | 12 | 8.32 | 260.33 | N.S. | 20 | 7.99 | 276.50 | N.S. |
| Random Grouped Slow (Dist. R) | 34 | 7.37 | 8.25 | | 44 | 8.33 | 264.38 | | 46 | 8.10 | 274.13 | |

[a] CAT Grade Placement Score

[b] STEP Converted Score

pupils of comparable ability in either mathematics or science achievement.

### Third and Fourth Year's Achievement

Achievement during the period between the end of the tenth and the end of the twelfth grade was measured by two forms of the STEP mathematics and science tests. In mathematics, the District A superior sample had significantly higher final adjusted mean scores than comparable District R pupils. Ability grouped pupils at the average level also had a significantly higher adjusted final mean score in mathematics. No significant difference was found between slow pupils in the two grouping systems. Comparisons of science achievement for District A and District R pupils of similar ability revealed no significant differences over this two year period (see Table 4-19).

### Overall Results for the Four Years

Analysis of the overall achievement gains for the four years of the study revealed significant differences in final adjusted mathematics achievement means, favoring District A average and District A slow pupils as compared with District R pupils of similar ability. No significant differences were found in mathematics between superior pupils in the two districts. Comparisons of District A and District R pupils on final adjusted science achievement means yielded no significant differences (see Table 4-20).

## ACHIEVEMENT RESULTS FOR A SAMPLE OF SUPERIOR JUNIOR HIGH SCHOOL STUDENTS

### Introduction

During the 1958-59 school year, the first year in which ability grouping was used in District A, some pressure was brought to bear upon the District R administration to introduce ability grouping into that district. As a result of this, it was decided by District R personnel to carry out a limited study of ability grouping in mathematics and science at the junior high school level and to delay any decisions on ability grouping until the project covered in this report had been completed. It was agreed that no pupils who were already in the sample of the research project reported herein would be included in this rather limited experiment. The project director worked closely with District R personnel in developing this project in order that the project would interfere as little as possible with the main study and also in order to provide a research design that would yield reliable information. This study was to be limited to seventh grade pupils of superior ability as defined by a measure of scholastic aptitude. Pupils selected for the study were in the top 15 per cent oñ the Lorge-Thorndyke Intelligence Tests. At the start of the 1959-60 school year 87 such pupils were placed in special accelerated classes in mathematics and science: one class in each of three junior high schools. Fifty-five pupils of comparable ability from these same schools who had been selected as members of Sample VII in the main research project were identified as a control group. It should be noted that placement of these superior pupils in special classes reduced the number of superior pupils in the regular heterogeneously grouped classes and to this degree had some effect upon the heterogeneity of Sample VII classes in District R. Pupils who were placed in the special classes will be referred to as the experimental group while comparable pupils in the heterogeneous mathematics and

Table 4-19

Achievement of Sample IX Pupils During Their Eleventh and Twelfth Grade Years

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean[a] | Final Mean | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 23 | 284.74 | 289.70 | 5.51* | 20 | 290.30 | 294.85 | N.S. |
| Random Grouped Superior (Dist. R) | 67 | 286.12 | 285.51 | | 47 | 294.00 | 295.02 | |
| Ability Grouped Average (Dist. A) | 68 | 270.66 | 277.57 | 6.57* | 62 | 281.87 | 285.56 | N.S. |
| Random Grouped Average (Dist. R) | 26 | 270.15 | 271.27 | | 11 | 281.90 | 284.91 | |
| Ability Grouped Slow (Dist. A) | 12 | 260.41 | 272.83 | N.S. | 18 | 280.16 | 278.33 | N.S. |
| Random Grouped Slow (Dist. R) | 24 | 252.92 | 263.29 | | 57 | 272.08 | 275.10 | |

[a] *STEP* Converted Scores

Table 4-20

Achievement of Sample IX Pupils Over a Four-Year Period

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean[a] | Final Mean[b] | $F_{y.x}$ | N | Initial Mean | Final Mean | $F_{y.x}$ |
| Ability Grouped Superior (Dist. A) | 22 | 10.22[a] | 287.32 | N.S. | 25 | 9.90 | 291.72 | N.S. |
| Random Grouped Superior (Dist. R) | 37 | 10.75 | 288.14 | | 37 | 10.34 | 289.22 | |
| Ability Grouped Average (Dist. A) | 36 | 8.37 | 279.22 | 6.14* | 35 | 8.34 | 287.23 | N.S. |
| Random Grouped Average (Dist. R) | 40 | 8.61 | 275.32 | | 40 | 8.79 | 283.97 | |
| Ability Grouped Slow (Dist. A) | 30 | 7.50 | 276.33 | 7.78 | 28 | 7.36 | 281.54 | N.S. |
| Random Grouped Slow (Dist. R) | 33 | 7.07 | 266.09 | | 32 | 7.37 | 275.78 | |

[a] CAT Grade Placement Scores
[b] STEP Converted Scores

science classes will be referred to as the control group. As a rule, the same teacher taught both homogeneous and heterogeneous classes in each school.

## Procedure

In this study the main curricular adjustment made for the homogeneous groups was an acceleration of the program, thus following the pattern set by District A. In mathematics, for example, the accelerated classes covered seventh and eighth grade mathematics, elementary algebra and plain geometry in three years. Pupils in heterogeneous groups completed seventh and eighth grade mathematics and elementary algebra but did not cover plain geometry. As the regular District R program for adjusting the curriculum to individual differences was based on enrichment in heterogeneous classes, superior pupils in the control group did participate in the enrichment program. Class size for the homogeneous and heterogeneous classes was closely comparable with a mean of 29 pupils per class. Class size stayed essentially the same although a few cases were lost and a few new pupils were added during the three year period. Complete data were collected for the three years on 78 of the 87 ability grouped pupils and 45 of the 55 random grouped pupils.

The STEP mathematics and science tests, form 3A were administered at the beginning of the study. Form 3B was administered at the end of the seventh grade. Form 2B, the high school form, was administered alternately with the Form 2A in the subsequent testing. Tests were administered at the beginning and end of the seventh, eighth, and ninth grades. The high school level tests were administered after the first year in order to provide a higher ceiling for measuring the achievement of these superior pupils.

## Results

Analysis of the results for each year was carried out using analysis of covariance in which the final mean achievement scores of the ability grouped and random grouped pupils were adjusted for differences in initial means. The $F_x$ values given in Table 4-21 are based on analysis at variance of the initial scores. The $F_{y.x}$ values are based upon the adjusted final means obtained in the analysis of covariance.

### Mathematics Achievement

It will be noted that on the initial STEP mathematics test administered at the beginning of the seventh grade the two groups were almost identical in mathematics achievement. The final adjusted means for the first year differ by about one point and indicate no significant difference. For each of the three years of the study the adjusted final means were not significantly different for the ability grouped and random grouped superior pupils. An F value of 3.92 would be required for significance at the .05 level and it may be seen that none of the yearly F tests on the adjusted end-of-year means approached this figure. Perhaps even more noteworthy was the extremely small difference found when the mathematics achievement gains from the beginning of the seventh grade to the end of the ninth grade were compared. The ninth grade means, when adjusted for seventh grade differences reflected the cumulative effects of three years in either heterogeneous or homogeneous classes. It will be noted that the F tests on

Table 4-21

Mathematics and Science Achievement of Superior Junior High School Pupils in Ability Grouped and Random Grouped Classes

| | | First Year | | | Second Year | | | Third Year | | | For Three Years | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N | Initial Mean | Final Mean | F | Initial Mean | Final Mean | F | Initial Mean | Final Mean | F | Initial Mean | Final Mean | F |
| **Mathematics** | | | | | | | | | | | | | |
| Ability Grouped | 78 | 269.29 | 277.51 | $F_x$ .00[a] | 274.29 | 285.66 | $F_x$ 4.75 | 287.68 | 289.37 | $F_x$ 1.13 | 270.51 | 289.13 | $F_x$ .46 |
| Random Grouped | 45 | 269.33 | 276.45 | $F_{x.y}$[b] 1.05 | 278.86 | 284.97 | $F_{x.y}$ 2.03 | 286.07 | 290.31 | $F_{x.y}$ 2.68 | 269.47 | 286.68 | $F_{x.y}$ .70 |
| **Science** | | | | | | | | | | | | | |
| Ability Grouped | 78 | 276.94 | 286.57 | $F_x$ .87 | 286.82 | 290.49 | $F_x$ 4.21 | 294.51 | 296.08 | $F_x$ 2.68 | 277.91 | 295.67 | $F_x$ .20 |
| Random Grouped | 45 | 278.27 | 283.38 | $F_{x.y}$ 8.72[c] | 283.40 | 288.42 | $F_{x.y}$ .00 | 291.47 | 292.04 | $F_{x.y}$ .65 | 277.89 | 293.08 | $F_{x.y}$ 2.17 |

[a] F-test on initial mean differences

[b] F-test on final adjusted mean differences

[c] Significant at .01 level, favoring ability grouped sample

these final adjusted ninth grade means was only .70 and the difference between the cumulative achievement gains of the two groups over this three year period was less than one and one-half points on the STEP converted score. These cumulative results agree closely with the Sample VII data previously discussed and lend strong support to the conclusion that ability grouping coupled with acceleration in junior high school mathematics offers no real achievement advantage for superior students as compared with random grouping coupled with enrichment.

## Science Results

The science results were generally in line with the findings reported in mathematics One significant difference however, was found in the yearly comparisons of science gains. It will be noted in Table 4-21 that the random grouped superior pupils started out slightly higher on initial achievement than the ability grouped sample and were somewhat lower at the end of the first year. The F test of the adjusted final means for the first year was significant beyond the .01 level in favor of the ability grouped pupils. This difference, however, had disappeared for the most part by the end of the second year. The third year differences were very small, yielding an adjusted F test of only .65. When science achievement over the entire three years of the study is compared it will be noted that the ability grouped pupils made a cumulative gain about two and one-half points greater on the STEP converted score than did the random grouped pupils. This difference, however, falls far short of significance at the .05 level.

## Discussion

In view of the fact that the ability grouped pupils were in an accelerated program in both mathematics and science and therefore had covered materials that had not been included in the curriculum for the heterogeneous classes, it is rather surprising that no achievement differences emerged. The lack of differences in science can probably be accounted for to a considerable degree by the enrichment program that permitted pupils to explore independently some of the areas that were covered as part of the accelerated curriculum in the ability grouped classes. Another factor to be considered in evaluating these results is that the STEP achievement battery is designed to measure basic concepts to a greater degree than most achievement tests and conversely measures less information of a highly specific nature. Such a measure of basic concepts and broad principles has much to recommend it, but it must be recognized that if a test emphasizing specific operations had been used, the homogeneous classes may have made higher scores.

Another possibility that might have helped explain the lack of differences in achievement scores was that the ability grouped classes, although progressing further in the subject area in their accelerated program had learned the material they covered less well. If this had been the case, the accelerated pupils would have made more errors than random grouped pupils on the less advanced problems because they had mastered this material less well; but would have made fewer errors on the more advanced problems which dealt with principles and concepts not yet reached by the heterogeneous classes. This possibility could best be tested through an item analysis of the final STEP mathematics test administered at the end of the ninth grade, as such an effect would be cumulative, thus showing itself most clearly on the final measure. It will be recalled that the mean scores of the two groups on the initial measure were very similar. Ten of the most advanced items on form 2B of the STEP test were identified and the percentage of pupils in each group who missed each item was computed. The results showed that an average of 39.8 per cent of the random grouped pupils missed these items as compared with an average of 40.9 per cent of the ability grouped pupils. Thus, we see that there was no substantial difference in the performance of the two groups on the advanced items of the test. These results lead us to reject the idea that the accelerated pupils had a better mastery of advanced materials.

## SUMMARY AND CONCLUSIONS

### Introduction

In this chapter we have examined in some detail the achievement data collected on the five samples of pupils studied in this research during the four years of the project. Let us now attempt to pull these data together in order to appraise the effects of ability grouping and random grouping upon achievement in the three major school divisions--elementary, junior high, and senior high school; and for the three ability levels--superior, average, and slow.

### Elementary School Achievement

Elementary school achievement data in this research are limited to the performance of Sample IV in grades four, five, and six, and Sample VI in grade six. A number of achievement areas as well as overall achievement were compared during the elementary school years for these samples. In studying Table 4-22, we find that during the elementary school years a total of 48 statistical comparisons were made between District A and District R pupils. Of these 48 comparisons, 24 were statistically significant. Eighteen of these significant differences favored District A pupils, while six favored District R pupils. At first glance, this would seem to indicate a rather marked advantage for ability grouping at the elementary school level. On looking further, however, we find that of the 18 significant differences favoring District A pupils, 15 occurred during the first year of the study. During the second year of the study, we find that all six of the statistically significant differences at the elementary level favor District R. These differences generally cancel out the apparent superiority of the District A pupils during the first year. When Sample IV pupils reached the sixth grade, during the third year of the study, only three significant differences occurred; all in science, and all favoring District A.

It is interesting to compare the sixth grade results obtained by Sample IV and by Sample VI, and speculate about the reasons for the much larger differences found among Sample VI pupils in the sixth grade, as compared with Sample IV pupils in the sixth grade. Several possible reasons for these differences have been previously discussed. During the pilot study year, when most of the elementary achievement differences favoring District A were found, we had the greatest possibility of Hawthorne Effect operating in favor of the District A program that was new to the district during that year. There is also a possibility that the use of the California Achievement Test Battery, which is the regular achievement battery for the districts, influenced the first year results. If these first year differences were due primarily to a true superiority of ability grouping over random grouping, we would expect the differences to grow larger each year as the cumulative effects of the better system widened the achievement

## Table 4-22

### Summary of Achievement Differences Between District A and District R Samples*

| | | Sample IV | | | | | | Sample VI | | | | Sample VII | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1st Yr. | 2nd Yr. | 3rd Yr. | 4th Yr. | 4 Yrs. | Last 3 Yr. | 1st Yr. | 2nd Yr. | 3&4 Yrs. | 4 Yrs. | 1st Yr. | 2nd Yr. | 3&4 Yrs. | 4 Yrs. |
| SUPERIOR | Math. | A[a] | NS | NS | A | | | A | R | NS | R | NS | A | NS | NS |
| | Lang. | A | | | | | | A | | | | | | | |
| | Science | | | A | NS | | | | NS | NS | NS | | A | NS | NS |
| | Reading | A | R | NS | | | | A | | | | | | | |
| | Soc. St | | | NS | | | | | | | | | | | |
| | Overall | A | NS | NS | | A | NS | A | | | | | | | |
| AVERAGE | Math | NS | R | NS | NS | | | A | NS | NS | NS | R | NS | NS | NS |
| | Lang. | A | | | | | | NS | | | | | | | |
| | Science | | | A | NS | | | | NS | A | A | | A | NS | NS |
| | Reading | A | R | NS | | | | A | | | | | | | |
| | Soc. St | | | NS | | | | | | | | | | | |
| | Overall | A | R | NS | | NS | R | A | | | | | | | |
| SLOW | Math. | NS | NS | NS | NS | | | NS | R | NS | NS | R | NS | NS | NS |
| | Lang. | NS | | | | | | A | | | | | | | |
| | Science | | | A | NS | | | | NS | NS | NS | | NS | R | NS |
| | Reading | NS | R | NS | | | | NS | | | | | | | |
| | Soc. St | | | NS | | | | | | | | | | | |
| | Overall | NS | R | NS | | R | R | NS | | | | | | | |

[a] A - District A sample made significantly greater achievement gains.
R - District R sample made significantly greater achievement gains.
NS - Difference between Districts A and R not significant.

| Sample VIII | | | | Sample IX | | | |
|---|---|---|---|---|---|---|---|
| 1st Yr. | 2nd Yr. | 3&4 Yrs. | 4 Yrs. | 1st Yr. | 2nd Yr. | 3&4 Yrs. | 4 Yrs. |
| NS | A | NS | NS | NS | NS | A | NS |
| | NS | NS | NS | | NS | NS | NS |
| R | NS | NS | NS | NS | NS | A | A |
| | A | NS | NS | | NS | NS | NS |
| NS | A | NS | NS | NS | NS | NS | A |
| | NS | NS | NS | | NS | NS | NS |

gap. As may be seen in Table 4-22, no such cumulative effects occurred, and in fact, most of the achievement differences favoring District A had disappeared by the time the Sample IV pupils had completed the sixth grade. These data clearly illustrate the dangers of drawing conclusions on the basis of a single year's research results.

## Junior High School Achievement

When we combine the mathematics achievement data for Sample IV, VI, VII, VIII, and IX, for the years when these pupils were in the junior high school, we find a total of seven one-year comparisons, one two-year comparison, and one four-year comparison at each ability level that reflect achievement gains made wholly or predominently during the junior high school years. Thus, combining the three ability levels for these samples gives a total of 27 statistical comparisons between ability grouped and random grouped pupils in mathematics during the junior high school years. It will be seen in Table 4-22 that of these 27 comparisons, four yielded significant differences in favor of District A, seven favored District R, and the remaining 16 were non-significan The predominance of non-significant differences would suggest that the grouping system had little effect upon mathematic achievement at the junior high school level. If any true difference exists, however, the larger number of statistically significant differences favoring District R would suggest that random grouping may be more effective in mathematics at this level. Two of the significant differences favoring District A were in grade nine. It will be recalled from the discussion of curriculum differences in Chapter 3 that, although ninth grade mathematics requirements are the same in two districts, a larger percentage of District A pupils take mathematics in the ninth grade. When this factor is considered the difference favoring District R appears somewhat more important than at first glance.

A total of six statistical comparisons between District A and District R pupils were made in junior high school science achievement at each ability level. These comparisons include four one-year comparisons, one two-year comparison, and one four-year comparison taken from Samples IV, VI, VII, and VIII. Of the total of 18 comparisons, five favored District A, and the remaining 13 were non-significant. The reader will recall however, that District A has a one-year science requirement at the eighth grade level, as compared with a half-year requirement in District R. It will be noted in Table 4-22 that of the five significant treatment differences in science achievement, three occurred at the eighth grade level. It is quite probable that these differences were caused by the difference in science requirements in District A rather than by the ability grouping system. As the remaining significant differences favoring District A occurred after the eighth grade, it is possible that this difference in eighth grade science requirements also had an influence upon these later differences. When all differences between ability grouped and random grouped samples in junior high school science achievement are considered, it would seem that the most tenable conclusion to draw from our data is that the grouping systems employed in District A and District R have little or no effect upon the science achievement of junior high school students.

## Senior High School Achievement

When we pull together all statistical comparisons between Sample VII, Sample VIII, and Sample IX pupils during their high school years, we find that a total of 42 comparisons were made. This includes all one-year comparisons at the high school level,

3

two-year comparisons in which both years or the second year were at the high school level, and four-year comparisons for the above samples. Of the 21 statistical comparisons of District A and District R pupils in mathematics achievement, 17 are non-significant, and four are significant favoring District A. (Thus, at the high school level it appears that ability grouping with acceleration may provide some advantage in mathematics over random grouping with enrichment.) All of the four significant differences for the high school years are found in Sample IX, and two of these cover cumulative achievement over the entire four years of the study. This suggests that ability grouping at the high school level may lead to some real advantages in mathematics achievement, although these advantages are small. On the other hand, these Sample IX differences may have been brought about by variables unrelated to ability grouping.

The same 21 comparisons in high school science achievement produced only one significant difference, favoring District R, with the other 20 differences being non-significant. The presence of a single significant difference in 21 comparisons would lead to the conclusion that the two groupings systems employed in the Utah study have little of no differential effect upon high school science achievement. The one significant difference favoring District R is probably the result of chance. In reviewing the high school data summarized in Table 4-22, the reader should recall that the high school data can be accepted with less confidence than either the elementary or junior high school data for the reasons previously discussed in Chapter 3.

## Achievement Summary by Ability Level

In comparing the achievement of District A and District R pupils within each of the three ability level classifications, we find that for each ability level a total of 47 statistical comparisons were made between comparable pupils in the five samples during the four years of the study. In reviewing the 47 comparisons between superior pupils in the two districts, we find 15 comparisons favoring District A, three favoring District R, and the remainder not significant. Of the 15 significant differences favoring District A, eight are found in the first year's results which may be less reliable than later data. The evidence, however, appears to favor ability grouping for superior pupils even if these first year differences are disregarded.

Examination of the 47 statistical comparisons between average pupils, reveal 13 favoring District A, six favoring District R, with the remainder not significant. These data would seem to suggest a slight achievement advantage for average pupils in the ability grouped classroom. When the first year's achievement results are omitted, however, the differences between the two treatments are somewhat reduced, although still favoring the ability grouped sample.

Comparisons between slow pupils in the two districts reveal four significant differences favoring District A pupils, seven favoring District R pupils, with the remaining 36 differences being non-significant. These data would seem to suggest that in terms of achievement the two grouping methods are about equally effective for slow pupils. If any advantage is present, it would appear from the data of this research that the random grouping system is slightly more effective for the slow pupil.

When all achievement data in this project are examined one must conclude that there is very little to choose from in terms of achievement between ability grouping with acceleration and random grouping with enrichment. The few differences that were found tended more often to favor the ability grouped samples, but, for the great majority of comparisons made, no significant differences were found. The statistically significant differences found were not large enough to suggest more than a slight advantage for one grouping system or the other with these advantages being far from consistent from one grade level or ability level to another. *Therefore, it is our conclusion that the decision to employ ability grouping or random grouping must be based upon considerations other than achievement*. The research herein reported considers a great many other variables such as self-concept, study habits, pupil attitudes, sociometric choice patterns, and personality that will help shed light on the overall advantages and disadvantages of these two grouping systems.

## REFERENCES

Billett, R. O. A controlled experiment to determine the advantages of homogeneous grouping. *Educ. Res. Bull.*, 1928, 8, 133-140.

Billett, R. O. *Provisions for individual differences, marking and promotion*. Bulletin 1932. Washington, D.C.: U.S. Office of Education, 1932, No. 17, Parts I, II, and III.

Bonar, H. S. Ability grouping in the first grade. *Elem. Sch. J.*, 1929, 29, [illegible]3-706.

Burtt, H. E., Chassel, L. M., & Hatch, E. M. Efficiency of instruction in unselected and selected sections in elementary psychology. *J. educ. Psychol.*, 1923, 14, 154-161.

Cook, R. R. A study of the results of homogeneous grouping of homogeneous grouping of ability in high school classes. *Yearb. Nat. Soc. Stud. Educ.*, 1924, 23.

Daniels, J. C. The effects of streaming in the primary school. *Brit. J. educ. Psychol.*, 1961, 31, 69-78.

Drews, E. M. *The effectiveness of homogeneous and heterogeneous ability grouping in ninth grade English classes with slow, average, and superior students*. Unpublished manuscript. Michigan State University, 1962.

Edmiston, R. W., & Benfer, J. C. Relationship between group achievement and range of abilities within groups. *J. educ. Res.*, 1949, 42, 547-548.

Ekstrom, Ruth B. *Experimental studies of homogeneous grouping a review of the literature*. Princeton, N. J.: Educational Testing Service, 1959.

French, J. W. Evidence from school records on the effectiveness of ability grouping. *J. educ. Res.*, 1960, 54, 83-91.

Goldberg, Miriam L., Justman, J. A., Passow, A. H., & Hage, G. *The effects of ability grouping*. Unpublished manuscript. New York: Teachers College, Columbia University, 1961.

Hansen, C. F. *A third year evaluation report on the four track curriculum*. Unpublished manuscript. Board of Education, District of Columbia, April 15, 1959.

Hartill, R. W. *Homogeneous grouping*. New York: Bureau of Publications, Teachers College, Columbus University, 1936.

Herr, W. Junior high school accelerates and their peers in senior high school. *Sch. Rev.*, 1937, 45, 186-195, 289-299.

Jones, V. A., & MacCall, W. A. Application of two techniques in evaluation some policies of dealing with bright children. *Teach. Coll. Rec.*, 1926, 27, 825-835.

Justman, J. Academic achievement of intellectually gifted accelerants and non-accelerants in junior high school. *Sch. Rev.*, 1954, 62, 142-150.

Kraraceus, W. C., & Wiles, E. An experiment in grouping for effective learning. *Elem. Sch. J.*, 1938, 3, 264-68.

Mann, H. How real are friendships of gifted and typical children in a program of partial segregation. *Except. Children*, 1957, 23, 199-201.

Martin, W. H. *The results of homogeneous grouping in the junior high school*. Doctoral thesis, Yale University, 1927.

Moyer, E. L. A study of the effects of classification by intelligence tests. *Yearb. Nat. Soc. Stud. Educ.*, 1924, 23, Part I.

Passow, A. H., Goldberg, M. L. & Link, F. R. Enriched mathematics for gifted junior high school students. *Educ. Leadership*, 1961, 18, 442-448.

Provust, M. M. Ability grouping in arithmetic. *Elem. Sch. J.*, 1960, 70, 391-398.

Purdom, T. L. *A scientific study to determine the value of homogeneous grouping made on the basis of intelligence tests*. Unpublished doctoral dissertation, University of Michigan, 1924.

Rankin, P. Evaluation of ability grouping in the seventh grade. *Rev. Educ. Res.*, 1931, 1, 223-225.

Rankin, P. T., Anderson, C. T., & Bergman, W. G. Ability grouping in the Detroit individualization experiment. *Yearb. Nat. Soc. Stud. Educ.*, 1936, 35, Part I.

Roberts, I. L. Homogeneous grouping: experiment. *Chicago Sch. J.*, 1947, 29, 30-32.

Simpson, R. E., & Martinson, R. A. *Educational programs for gifted pupils*. Sacramento, California: California State Department of Education., January, 1961.

Ullrich, O. A. *An experimental study of the effect on learning of sectioning college classes on the basis of ability*. Unpublished doctoral dissertation, University of Texas, 1926.

West, J., & Sievers, C. Experiment in cross-grouping. J. educ. Res., 1960, 54, 70-72.

Westmayer, P. Just stand out of the way! Sch. Sci. Math., 1957, 57, 643, 646.

Woody, C. Results obtained from ability grouping. Bureau Educ. Ref. Res. Bull., 1924.

Worlton, J. T. The why of homogeneous classification. Elem. Sch. J., 1926, 27, 265-274.

Worlton, J. T. The effect of homogeneous classification of scholastic achievement of bright pupils. Elem. Sch. J., 1928, 28, 336-345.

## Chapter 5

# REALIZATION OF ACHIEVEMENT POTENTIAL

by Walter R. Borg and M. Nephi Manning[1]

## INTRODUCTION

Although a number of studies have been conducted in which achievement under different grouping systems has been compared, achievement of individual pupils in these systems as related to the achievement one might expect on the basis of the individual's ability has not been explored. In this phase of the Utah study, samples of elementary school children from random grouped and ability grouped classrooms were classified as overachievers, underachievers and normal achievers based on the relationship between their achievement, chronological age, and mental age. The proportions of overachievers, normal achievers and underachievers among pupils of different ability levels in the two systems were then studied in an effort to obtain insight into the question of who is motivated to achieve under these two methods of grouping. In the second phase of this study, descriptive data were collected on samples of underachievers and normal achievers in an attempt to identify possible causes of underachievement.

This is no general agreement among educators on a precise definition of "overachiever" and "underachiever." Most definitions, however, provide for a comparison between the individual's achievement and his achievement potential as estimated by an intelligence test. Underachievers are defined as achieving at some point below their potential achievement. Overachievers are less frequently discussed. Most educators have difficulty accepting the concept of the "overachiever" as this implies that the individual is achieveing more than his potential. In this study we have defined the underachiever as a person who is achieving more than nine months below the normal level expected on the basis of mental age and chronological age, and the overachiever as a person achieving the same amount above this expected level. The expected level, however, should not be thought of as the individual's maximum potential. The concept of the overachiever becomes more acceptable if we consider the following analogy. Let us say that the normal achiever is similar to the average automobile that has eight cylinders but generally operates on only five or six of them. The underachiever then becomes comparable to the auto that lurches along unsteadily on three or four of its eight cylinders while the overachiever may be compared with that rare vehicle that speeds along on seven or perhaps even all eight cylinders. Thus, the overachiever performs better than the average achiever of comparable ability, and more closely approaches his achievement potential.

---

[1] A portion of this phase of the Utah study is based on Dr. Manning's dissertation that was completed under Dr. Borg's supervision.

## WORKING HYPOTHESES

The following working hypotheses were established with respect to the performance of overachievers, underachievers and normal achievers in ability grouped and random grouped classrooms:

1. There will be no significant difference in the overall proportion of overachievers, underachievers and normal achievers in the ability grouped sample as compared with the random grouped sample.

2. There will be no significant differences in the proportions of overachievers, normal achievers and underachievers between the fourth and sixth grade samples.

3. There will be significantly more boys than girls who are classified as underachievers and fewer boys who are classified as overachievers.

4. There will be a significant tendency for superior pupils in the ability grouped system to be more frequently overachievers and less frequently underachievers than in the random grouped system. This hypothesis is based on the observation that schools employing random grouping generally slant the level of work to the average pupil and in so doing may fail to challenge the superior pupil.

5. There will be no significant difference between the ability grouped system and the random grouped system in the proportion of overachievers, normal achievers and underachievers among average pupils. Because of the tendency to aim the curriculum at the average pupil in heterogeneous classes, the two grouping treatments involved in the Utah study would be most similar for pupils at this ability level. One difference between the treatments that may affect the average pupil, is his association with superior pupils in the heterogeneous class. Some educators have maintained that the average pupil is motivated to greater achievement because of the presence of superior pupils in the heterogeneous classroom. If this claim is valid, more overachievers and fewer underachievers should be found in the District R average sample.

6. There will be significantly fewer underachievers and more overachievers among slow pupils in the random grouped system than in the ability grouped system. This hypothesis is based on the same rationale as hypothesis four. It is anticipated that random grouping, by slanting instruction to the average pupil, will tend to put pressure on the slow pupil to achieve.

The second phase of the study was concerned primarily with comparisons between underachievers and normal achievers. These groups were compared on several dependent variables. The working hypotheses established in this phase of the study are:

7. Underachievers will obtain significantly less favorable scores than normal achievers in measures of study habits, emotional disturbance, achievement motivation, socioeconomic status and parents' attitudes towards school.

8. There will be no significant differences between underachievers and normal achievers in incidence of health problems.

## PREVIOUS RESEARCH RELATED TO UNDERACHIEVEMENT

### Introduction

The studies reviewed have been divided into those dealing with environmental factors that might relate to underachievement such as family background, home conditions, social position of parents, community and neighborhood environment, companions and age preferences, counseling procedures, teacher influence, and social acceptance--and into those dealing with personal factors that might relate to underachievement such as, study methods, aspirations, self-concept, emotional disturbance, health, intelligence quotient, goals and achievement motives, interests, values, and occupational choices.

## Environmental Factors

In evaluating the influence home conditions have upon underachievement, it was found that parents' attitudes toward school and socioeconomic conditions in the family are major factors dealt with in the literature. The work of Goldberg and her associates (1959) pointed out that home environment is a contributing factor to underachievement. In this experiment, a classroom climate was maintained where the students felt free to come to the teacher with personal, as well as with school problems. These personal contacts revealed that in many instances there were serious misunderstandings and poor relationships between boys and their parents that were related to school behavior.

A study by Gibboney (1959), was concerned with socioeconomic status and achievement in social studies. The investigator compared the achievement of two groups of sixth grade students: one from the upper-middle class and one from the upper-lower class. Differences in intelligence were controlled by covariance analysis. He found that the group of higher status achieved significantly more (.05 level).

Williams (1959), also sought to determine if social acceptance significantly influences performance. She administered the California Test of Mental Maturity, Classroom Social Distance Scale and other measures to a total of 888 children. Analysis of the data revealed that four out of five children who were high in total acceptance were achieving within or beyond expectance. She summarizes as follows:

1. No appreciable differences in intelligence between high and low acceptees.
2. An appreciable difference in both test-measurements and scholastic evaluation of social performance favoring the high acceptees.
3. Evidence of greater acceptance extended to the group by the high acceptees.
4. No difference between high and low acceptees with regard to characteristics valued in friends.
5. Greater permissiveness and wider variety of channels for learning in the experience of the high acceptees.
6. More consistent satisfaction with the inter-personal relationships among the high acceptees, and more consistent dissatisfaction among the low acceptees. (Williams, 1959, p. 43)

Armstrong (1955) compared the interests and social adjustment of underachievers and normal achievers at the secondary school level. The procedure followed was to administer the Otis Quick-Scoring Mental Ability Test to all students in grades nine and eleven in Woodrow Wilson High School, Middletown, Connecticut to establish a common standard of intelligence. Otis I.Q. scores and the average of school marks were correlated. Those whose average school mark was at least one-half of one standard error of prediction below their predicted average were considered underachievers. These pupils were matched on age, sex, grade, and curriculum with a control group of normal achievers whose average school marks were within one-half of one standrad error of prediction from the predicted average. Kuder Preference Records, were administered to all students in the study. Scores above the 75th percentile and below the 25th percentile were compared for underachievers and normal achievers. Rating scales on "cooperation," "dependability," and "judgment" were compeledon all students by selected teachers. Each student was interviewed by a counselor and personal records were used to clarify data obtained in the interviews. Chi square was used as a test of significance for all comparisons between underachievers and normal achievers. The results indicated that underachievers were more likely than normal achievers--

1. To be following vocational goals set for them by others.
2. To have stated goals not in line with their dominant interests.
3. To prefer outdoor activity.
4. To have been chosen less often for positions of responsibility (girls only).
5. To prefer companions older than themselves (boys only). (Armstrong, 1955, p. 1349)

## Personal Factors

The research previously cited by Goldberg and her associates (1959) also analyzed the causes of underachievement in superior pupils. This study attempted to identify social and personal factors associated with underachievement, and to experiment with school procedures that would provide special attention to the problems of an underachieving group. Of 102 underachievers entering the tenth grade, 70 were paired on the basis of I.Q. and ninth-year averages. One student from each pair was placed in a special class and the other became a control. The control students were not identified to themselves or to their teachers and were randomly distributed in homeroom situations and subject matter classes. In addition, a group of high ability high achievers was identified. Records were kept on these students, but no special provisions were made for them beyond those normally made by the school for able students.

A testing and interview program was designed to bring out the differences between these groups. The Iowa Tests of Educational Development were administered as a regular part of the school's testing program. Also, objective measures of self-attitudes, attitudes toward school, family patterns, problem areas, academic aspiration levels, and vocational choices were obtained from each student in the study.

The achievement test results indicated that achievement on objective measures of academic mastery is more closely related to intelligence than to school grades. The attitude and personality measures provided insight into the lives and attitudes of bright young adolescents. It was found that disruption of the normal family pattern was much more frequently observed among the underachievers. High achievers were more satisfied with school and with their school performance than were the underachievers, and had higher aspiration levels.

The interview produced information in regard to friendships that was revealing. All of the underachievers took part in social activities and had friends. Frequently, however, they attributed their social success to "not being a grind," to not studying too hard or too much. They recognized the fact that they were potentially capable of outstanding academic achievement, but in most of the interviews there was strong resistance against making the necessary effort.

Winkler and MacNutt (1960) studied underachievement at the fourth grade level for the purpose of understanding, if possible, reasons why children failed to continue the growth of which they seemed capable during the first three years in school. Information was sought concerning underachievement as it related to intellect, personality, socioeconomic conditions, family relationships, health and developmental patterns, and peer and authority relationships. The subjects were 271 pupils (148 boys and 123 girls) in the fourth grades from eight schools in Tangipahoa Parish, Louisiana.

The *California Achievement Test Battery* was used to locate underachievers who were defined as those whose grade placement on two or more test areas was below 4.0 and whose level of intelligence was above 90 I.Q. Intelligence was measured using the *SRA Primary Mental Abilities*. The *SRA Junior Inventory*, and *Ohio Social Acceptance Scale* were also administered. Other data involving the family were obtained by a questionnaire administered to the parents.

Among the findings were that

> . . . girl underachievers have more personality problems than boy underachievers, apparently because it is a threat to a girl's self-concept. Underachievers were concentrated at the middle socioeconomic level, with a small percentage in the lower level and none at the upper middle class. Social class level, parental pressures, and level of aspiration are the significant factors in the child's underachievement. (Winkler and MacNutt, 1960, p. 58)

The self-concept factor has also been studied as a possible cause of underachievement. The results of part of a research project being performed by Brookover, Velinsky, and Thomas (1961), of Michigan State University, were made available to the investigators. This study attempts to learn if achievement is primarily related to the individual's self-conception of his ability irrespective of his measured I.Q. The research sample includes seventh grade pupils for whom two sets of I.Q. scores were available. On the basis of the two sets of I.Q. scores and school marks for the fourth through the sixth grades, a sample of 1,151 students was obtained. From this number, the investigators identified the overachievers and underachievers and eliminated the 365 students who fell in the middle range.

The results of this study demonstrated a positive relationship between achievement and self-concept. The higher the achievement scores, the higher were the pupil's self-concepts relative to ability. It was also noted that high achievers had a higher mean self-concept score than underachievers even though these two groups had comparable I.Q.'s. The level of significance of this difference was .001.

What part does emotional disturbance play in underachievement? Is it possible that conflicts, anxiety, and other manifestations of poor adjustment are caused by poor academic ability and in turn contribute to poor scholarship?

Jensen (1958) experimented with four groups of college students including non-achieving students of high and low ability, and achieving students of high and low ability. All students were administered the *Minnesota Multiphasic Personality Inventory* (*MMPI*). The most pronounced trend in this study was for the non-achieving students of low ability to consistently obtain less favorable scores on the *MMPI* than achieving students of high ability.

> These findings offer some support for the general hypothesis that students of low scholastic ability as compared to other groups are at a disadvantage with respect to non-intellectual areas of college life. With certain exceptions, there was a general tendency throughout the study for non-achievers of low scholastic ability to encounter more adjustment problems than other students with whom they were compared. Thus, these students tended to be at a disadvantage with respect to non-intellectual areas, as measured by the MMPI, as well as in their academic pursuits. The general trend was for gifted achievers to express themselves as having fewer adjustment problems than the other groups. The results presented throughout this study support the findings reported by Terman, Brown, Hinkelman, Lightfoot, and others. These studies, including this one, all have a tendency

> to show that scholastic ability favors adjustment, while low scholastic ability obstructs it. (Jensen, 1958, p. 500)

A project concerned with early identification of emotionally disturbed children was reported by Bower, Tashnovian, and Larson (1958). This project was undertaken by the California State Department of Education and involved 75 school districts, 200 teachers, and approximately 5,500 school children. There was at least one clincially designated emotionally disturbed child in each class studied. Altogether there were 162 boys (approximately 3 per cent) and 45 girls (approximately 1 per cent) in the group of emotionally disturbed children. Among the findings of this research, those most pertinent to the Utah study were:

> 1. The emotionally disturbed children scored significantly lower on group I.Q. tests. On psychological tests given individually, they approached the mean of all children included in the study.
>
> 2. The emotionally disturbed children scored significantly lower on reading and arithmetic achievement tests. The differences were greater and more significant on arithmetic achievement. The higher the school grade, the greater the differences between the emotionally disturbed child and others in the classes. (p. 18)

Stone and Ganung (1956) reported a four year study of 126 freshman students that was designed to determine if differences in scholastic achievement could be found between those who had high scores and those who had normal scores on the Minnesota Multiphasic Personality Inventory. They found that significantly more of the normal group graduated as compared to the high scoring group on the MMPI (38 per cent as opposed to 22 per cent). Girls who scored high on one or more scales of the MMPI also had a lower grade point average than those receiving normal scores.

Little attention has been given to the health and physical characteristics of underachievers. Klausmeier (1958) studied the physical, behavorial, and other characteristics of high and low achieving children in favored environments.

> The means of height, weight, strength of grip, permanent teeth, and carpal development of high achieving children were found to be not significantly different from the means for low achievers in third, fourth and fifth grade classes in favored environments. (Klausmeier, 1958, p. 580)

Some research has been concerned with the interests and motivation of underachievers. Armstrong (1955) studied the interests and social adjustment of underachievers as they compared to a matched group of normal achievers at the secondary school level. Statistically significant differences between underachievers and normal achievers were found as follows:

> Underachievers were found more often (1) to have chosen their future occupations because of the influence of others, (2) to have future vocational goals which did not agree with their dominant interest as measured by the Kuder Preference Record-Vocational, (3) to have obtained a greater number of low scores on the computational scale of the Kuder Preference-Vocational, (4) to have obtained more low scores in the area of "preference for avoiding conflicts" as measured by the Kuder Preference Record-Personal, (5) to have obtained lower ratings on "cooperation," "dependability," and "judgment," (6) to prefer companions older than themselves (boys only), (7) to have obtained a smaller number of high scores on the computational scale of the Kuder Preference Record-Vocational, and (8) not to have been chosen for positions of responsibility in extra-curricular activities (girls only). (Armstrong, 1955, p. 1349)

## Summary

Environmental factors relating to underachievement that have been mentioned in the literature include:

1. Lack good models in education and occupation in those they admire in home and community (Goldberg et.al., 1959; Russell, 1958).
2. Low educational level of family (Goldberg, et. al., 1959).
3. Unstable family structure (Goldberg, et. al., 1959).
4. Lower socioeconomic status (Hollinshead and Redlich, 1958; Gibboney, 1959; Martinson, 1960; and Stivers, 1959).
5. Ethnic and religious prejudice (Goldberg, et. al., 1959).
6. Rejection often by classmates for personal or caste reasons (Bedoian, 1954; Muggenthaler, 1955).
7. Often over-age in their peer group, causing rejection (Bedoian, 1954).
8. Environment too complicated for their ability (MacLean, 1958).
9. Have often received coercive influence from teachers or others (Rosenfield and Zander, 1961; Fliegler, 1957).
10. Suffer more pressures (Goodlad, 1960).
11. Blocks to learning have been set up (Goodlad, 1960).

Personal factors relation to low achievement and underachievement that have been mentioned in the literature include:

1. Underachievers dislike or fear people--especially parental or authoritative figures, siblings, and peers (Goldberg, et. al., 1959; Winkler and MacNutt, 1960).
2. Unable to create warm relationships with others (Burdick, 1961; Bower, Tashnovian, and Larson, 1958).
3. Have lower or unrealistic levels of aspiration (Winkler and MacNutt, 1960; Sears, 1940).
4. More often follow vocational goals set for them by others (Armstrong, 1955).
5. Have fewer vocational and other interests (Armstrong, 1955; MacLean, 1958).
6. Have stated goals not in line with their dominant interests (Armstrong, 1955; MacLean, 1958).
7. More prone to develop mental illness; more emotionally disturbed; have more difficulty in making non-academic adjustments (Jensen, 1958; Stone and Ganung, 1956; Ullman, 1952; Bower, Tashnovian, and Larson, 1958).
8. Have lower threshold of frustration (Fliegler, 1957).
9. Have lower intelligence quotient (Jensen, 1958; Hinkleman, 1955; Bower, Tashnovian, and Larson, 1958; Frandsen, and Higginson, 1951).
10. Have low self-concept of ability (Brookover, Velinsky, and Thomas, 1961; Goldberg, 1959; Bower, Tashnovian, and Larson, 1958).
11. Prefer to avoid conflicts (Armstrong, 1955).
12. Less cooperative, less dependable, have poorer judgment (Armstrong, 1955).
13. Prefer outdoor activity (Armstrong, 1955).
14. Prefer companions older than self (Armstrong, 1955).

# PROCEDURE

## Sample

The data collected in the Utah study that deals with achievement potential are based upon measures obtained on pupils in Samples IV and VI during the 1958-59 and 1959-60 school years. Three different sub-samples taken from Sample IV and VI were employed in this section of the study. The first phase of the study is concerned with comparisons between District A and District R on the numbers of overachievers, normal achievers, and underachievers. In this phase of the study, a total of 674 District A pupils and 793 District R pupils were compared. These pupils were in grades four and six and were classified in terms of achievement catagory (overachievers, normal achievers, and underachievers), sex, and ability level (superior, average, and slow). Table 5-1 provides a breakdown of the subjects in terms of the aforementioned categories.

The classification of pupils into the three ability levels had already been accomplished in connection with the achievement analysis. It will be recalled that cut-off points were established on the California Achievement Test Battery (CAT) which was administered to subjects at the beginning of the school year. These cut-off points made it possible to assign ability levels to District R pupils that were comparable to the ability levels established by the ability grouping procedure in District A.

In order to identify underachievers, the mean achievement score for fourth and sixth grade pupils on form W of the California Achievement Test Battery administered at the end of the school year was calculated for both districts. The overall mean grade placement score on the CAT for the District A sixth grade group was 7.58 as compared with an overall mean grade placement for the District R group of 7.48. Comparable means for the fourth grade were 5.67 and 5.68.

The classification of pupils as underachievers and overachievers presents some problems. A number of methods have been used from time to time including the rating of any pupils who are not achieving up to their normal grade placement as underachievers. Such a system ignores the pupil's achievement relative to his ability. Another system defines the underachiever as a pupil whose educational age is lower than his mental age. This system seems to place excessive weight on ability and tends to classify the majority of bright pupils as underachievers. Although the bright pupil may be expected to score well above his grade placement in subjects such as reading which are not closely tied to the learning of specific sequences of operations, it is doubtful whether it is reasonable to expect the bright pupil to score greatly above his grade placement in subjects such as mathematics where a knowledge of certain operations is essential for working the more advanced problems.

In the Utah study a system of classification for underachievers developed by the Santa Monica, California Unified School District was employed. The formula used in this case determines the expected grade placement by taking the sum of twice the pupil's mental age plus his chronological age and dividing by three. The mental age used in this study was based upon the Language Score of the California Test of Mental Maturity[1] (CMM). This formula is, of course, arbitrary, but seems to be more

---

[1] Because the CMM was administered earlier in the school year than the CAT, an adjustment in obtained mental age was made to compensate for the difference in time of administration of the two measures.

Table 5-1

Percentage of Each Subgroup Classified as Overachievers, Normal Achievers and Underachievers

| Subgroup | N | Percent Over-Achievers | Percent Normal-Achievers | Percent Under-Achievers |
|---|---|---|---|---|
| 4th Grade Sample | | | | |
| Superior Boys, District A | 44 | 40.9 | 54.5 | 4.6 |
| Superior Girls, District A | 62 | 51.6 | 46.8 | 1.6 |
| Superior Boys, District R | 81 | 23.4 | 64.2 | 12.4 |
| Superior Girls, District R | 96 | 38.5 | 57.3 | 4.2 |
| Average Boys, District A | 60 | 25.0 | 58.3 | 16.7 |
| Average Girls, District A | 41 | 41.5 | 51.2 | 7.3 |
| Average Boys, District R | 47 | 19.1 | 48.9 | 32.0 |
| Average Girls, District R | 45 | 35.6 | 48.8 | 15.6 |
| Slow Boys, District A | 53 | 28.3 | 52.8 | 18.9 |
| Slow Girls, District A | 23 | 39.1 | 39.1 | 21.8 |
| Slow Boys, District R | 35 | 11.4 | 54.3 | 34.3 |
| Slow Girls, District R | 20 | 25.0 | 65.0 | 10.0 |
| 6th Grade Sample | | | | |
| Superior Boys, District A | 65 | 23.3 | 52.8 | 13.9 |
| Superior Girls, District A | 73 | 49.3 | 45.2 | 5.5 |
| Superior Boys, District R | 116 | 11.2 | 63.8 | 25.0 |
| Superior Girls, District R | 123 | 14.6 | 68.3 | 17.1 |
| Average Boys, District A | 64 | 15.6 | 59.4 | 25.0 |
| Average Girls, District A | 66 | 18.2 | 66.7 | 15.1 |
| Average Boys, District R | 82 | 7.3 | 53.7 | 39.0 |
| Average Girls, District R | 67 | 25.4 | 62.7 | 11.9 |
| Slow Boys, District A | 76 | 5.3 | 59.2 | 35.5 |
| Slow Girls, District A | 47 | 14.9 | 68.7 | 17.0 |
| Slow Boys, District R | 55 | 9.2 | 45.4 | 45.4 |
| Slow Girls, District R | 26 | 3.8 | 57.7 | 38.5 |

reasonable than other systems that are being employed as it takes into account both mental age and chronological age in arriving at an expected educational age score. Each pupil in the sample was given a score that represented the difference between his expected educational age in months, based on the Santa Monica formula and his obtained educational age in months based on the California Achievement Test Battery, form W. This battery was found to yield educational age scores consistently higher than expected.[1] The median educational age on the CAT was four months higher than the median expected educational age calculated by the Santa Monica formula. This median was used as the center of the normal achieving group. Pupils whose education age was nine months or more above this median were classified as "overachievers," while those whose educational age was nine months or more below this median were classified as "underachievers."

The primary purpose of the second phase of the study was to collect comparative data for a sample of normal achievers and underachievers on a number of dependent variables. A stratified random sample taken from the Sample VI pupils used in the first phase of the study was employed. The same sample of underachievers were used for collecting data on all dependent variables in the second phase of the study. The normal achiever sample, however, was reduced somewhat for collection of data using the individual interview and physical examination because of the high cost of collecting these data. Table 5-2 shows the number of cases employed in the second phase of the study.

## Measures Used

### California Short-Form Test of Mental Maturity

The California Short-Form Test of Mental Maturity, 1957, S-Form (CMM) (for grades four through eight) was used in this phase of the Utah study, primarily in the identification of underachievers. This test samples mental processes in four areas: spatial relationships, logical reasoning, numerical reasoning, and verbal concepts. The primary purpose of the test was to obtain mental ages and language I.Q.'s needed for establishing the expected age-grade placement of students.

### California Achievement Test Battery (CAT)

This battery was used to obtain data on initial and end-of-year achievement. The initial battery was used to classify pupils into the three ability levels, and the end-of-year battery was used in the Santa Monica formula to classify pupils as under-achievers, normal achievers or overachievers. This battery yields scores in reading, mathematics, and language.

The reader will recall that there were a number of limitations in the first year's achievement data in the Utah study. These limitations were discussed in Chapter 4 and should be kept in mind in interpreting the data presented in this chapter because the first year's achievement data were employed to classify pupils as overachievers,

---

[1] Other phases of the study demonstrated that pupils in both districts consistently received average grade placement scores on the CAT form W that were higher than the pupil's grade placement.

Table 5-2

Samples of Underachievers and Normal Achievers used in the 1959-60 Phase of the Study

For data on study methods, emotional disturbance and achievement motives:

| | District A | | District R | |
|---|---|---|---|---|
| | Boys | Girls | Boys | Girls |
| Underachievers: | | | | |
| Above average | 10 | 10 | 10 | 10 |
| Average | 10 | 10 | 10 | 10 |
| Below average | 10 | 10 | 10 | 10 |
| Total | 30 | 30 | 30 | 30 |
| Normal Achievers: | | | | |
| Above average | 15 | 15 | 15 | 15 |
| Average | 15 | 15 | 15 | 15 |
| Below average | 15 | 15 | 15 | 15 |
| Total | 45 | 45 | 45 | 45 |
| District Total | 150 | | 150 | |
| Total sample (both districts) | | 300 | | |

For individual interview data and physical examination:

| | District A | | District R | |
|---|---|---|---|---|
| | Boys | Girls | Boys | Girls |
| Underachievers: | | | | |
| Above average | 10 | 10 | 10 | 10 |
| Average | 10 | 10 | 10 | 10 |
| Below average | 10 | 10 | 10 | 10 |
| Total | 30 | 30 | 30 | 30 |
| Normal Achievers: | | | | |
| Above average | 8 | 8 | 8 | 8 |
| Average | 8 | 8 | 8 | 8 |
| Below average | 8 | 8 | 8 | 8 |
| Total | 24 | 24 | 24 | 24 |
| District Totals | 108 | | 108 | |
| Total sample (both districts) | | 216 | | |

normal achievers, and underachievers in this phase of the study. Pupils in ability grouped classrooms generally made much greater achievement gains during the first year of the study than comparable pupils in random grouped classrooms. This trend, however, was not maintained during the subsequent years of the study. These greater first year gains, however, were reflected in the classification of pupils as overachievers normal achievers, or underachievers in this phase of the study. This would tend to increase the number of pupils classified as overachievers and decrease the number of pupils classified as underachievers in District A. Unfortunately, sufficient funds were not available to classify pupils in terms of their achievement potential for the remaining years of the study. Thus, a check of the first year's results is not available at this time. It is hoped that these data will be analyzed at a later date in order to provide further information on the degree to which pupils realize their achievement potential in the two grouping systems.

## California Study Methods Survey (CSMS)

This measure was designed for use in grades seven to thirteen and is a self-report inventory concerned with the study methods and attitudes of the student. The CSMS is made up of 150 standardized questions designed to yield four scores: Attitude Towards School, Mechanics of Study, Planning and System, and a Verification Score. The Verification Score empirically identifies students who do not give conscientious and valid answers. The total battery score was used in this phase of the Utah study. See Chapter 6 for further discussion on this test.

## Thinking About Yourself

The Thinking About Yourself (Bower et. al., 1958) inventory was used to measure the degree of emotional disturbance in students classified as normal achievers and underachievers. The development of this inventory was based upon the assumptions that if a child with a developing personality disorder can be identified early, he can be helped more effectively to remedy the condition causing the disorder, and that personality disorders in adulthood are the result of a progressively developing condition assessable from the child's behavior. It utilizes the intra-individual measure of two concepts--that of self, and that of wanted or desired self. "Self" is defined as that which the child can afford to reveal in his answers to the questions. The sums of the differences or discrepancies between "self" and "wanted self" was used to determine the amount of emotional disturbance among children who were classified as underachievers as compared with students who were classified as normal achievers. This measure differs from the Index of Adjustment and Values, the principal self-concept measure used in the Utah study, in a number of ways. The most important differences were that the Thinking About Yourself inventory used behavior descriptions instead of trait words, and is more directly related to the school situation.

After obtaining the discrepancy scores for each set of questions, the scores were summed to get the student's total discrepancy score. A large discrepancy score indicates a large divergence between the pupil's "self" and "wanted self," and is considered to be indicative of a greater degree of emotional disturbance. Validity evidence on this measure is limited at the present time.

### Thematic Apperception Test (TAT)

This well known projective technique is based on the assumption that in completing or structuring an incomplete or unstructured situation presented by means of a picture, the subject projects his own personality which will be revealed by an analysis of his responses. For the purposes of this study, four picture cards were used: 1, 2, 4, and 16, to obtain a measure of achievement motivation. Group administration was employed by flashing each picture on a screen and having the students write their responses. John (1960) found that these particular cards were the most productive in revealing achievement motives. Analysis of the stories was limited to achievement motives only, using the Atkinson (1958) scoring system for achievement motivation.

In order to measure scorer reliability, some of the cases were scored twice and product-moment correlation coefficients were computed to establish the relationship of the first to the second score. The reliability coefficients for two randomly selected sets of responses were .81 (on 60 cases) and .91 (on 36 cases).

## Other Measures

### Social Class

To obtain information of socioeconomic status and parents' attitudes, a questionnaire was developed which contained: a modification of Hollingshead's occupational and educational scales, used by the investigator to determine the socioeconomic status; questions concerned with the student's self perception; questions regarding home conditions and parent's attitude toward school. These questions were of such a nature that a score could be assigned so as to permit quantitative analysis.

Inasmuch as it has been shown that a high degree of reliability could be obtained by the use of the occupation and education of the chief wage earner alone, (Hollinshead and Redlich, 1958) a modified Hollingshead approach was used, based on these two factors. The Hollingshead educational and occupational seven-point scales were marked at the time of a personal interview with each student by the researcher. The final result placed the family into one of five social classes.

### Health

To collect health data on normal and underachievers, a case history form was filled out on each student in the sample by the City and County Nurses, and a health examination was given by a practicing physician. These measures were used to determine if differences in physical health was a significant factor in academic achievement. The directors of the City and County Health Center and the nurses assigned to the various schools met in conference with one of the investigators to determine which health variables would be measured and to plan the data collection procedures.

Clinics were set up in the schools to carry out the physical examinations, and school nurses visited the pupils' homes where past history data were obtained from the parents. A practicing pediatrician was employed to make the physical examinations, which were given at the respective schools and required about 15 minutes each.

Following the health examination, each student met with one of the investigators in private and responded orally to items on the questionnaire that related to achievement motives, self-concept, home conditions and parents' attitude toward school.

Their answers were assigned a quantitative score. The dichotomy of "health problem"--"no health problem" had as its line of demarcation three or more identified problems discovered at the time of the physical examination. The student who had fewer than three identified problems was classified as having "no health problem."

Parent's attitudes toward school, and the comparisons on the social class data were analyzed for significance using the t-test.

## RESULTS

### Introduction

The first phase of the study was concerned with differences in the proportions of overachievers, normal achievers, and underachievers in the two grouping treatments The significance of differences in the proportions of pupils in these three classifications were determined using chi square contingency tables. Overall comparisons between pupils in the two grouping treatments were made as well as comparisons between the sexes, between pupils in the fourth vs. sixth grades, and between District A and District R pupils at a given ability level.

### Sex Differences

The first comparisons that were made were aimed at determining whether sex differences occurred in the overall proportion of overachievers and underachievers within the two treatments. Chi square contingency tables were set up comparing boys with girls in the proportion of underachievers, normal achievers and unceråchievers. Examination of the chi square contingency tables indicated that in every case there was a greater proportion of overachievers and a lesser proportion of underachievers among the girls. Some of these differences, however, were not significantly different.[1] In District A, it was found that a difference significant at the .05 level occurred (chi square equaled 8.45) between fourth grade boys and girls. Fourteen per cent of the fourth grade boys in the sample were classified as underachievers as contrasted with 7.1 per cent of the girls. At the other end of the distribution, 46 per cent of the girls in this sample were classified as overachievers as compared with 30.6 per cent of the boys. A more pronounced trend existed between fourth grade boys and girls in District R. In this group, 22.7 per cent of the boys were classified as underachievers as contrasted with 8.1 per cent of the girls. Thirty six per cent of the fourth grade girls in this group were overachievers as compared to 19.6 per cent of the boys.

In the sixth grade samples, the tendency toward more overachieving girls and fewer overachieving boys was more marked than the fourth grade data in District A, but was closely comparable to the fourth grade results for District R.

After establishing the above overall sex differences in the proportions of overachievers and underachievers in each sample, similar comparisons were made for superior, average, and slow pupils within each group. In order to determine the ability levels at which these sex differences operated most strongly, the fourth and sixth grade samples were combined for this analysis to provide larger subgroups.

---

[1] Unless otherwise indicated, a difference reported as significant will be significant beyond the .01 level.

In the ability grouped sample, superior fourth and sixth grade girls were found to have a significantly greater proportion of overachievers and lesser proportion of underachievers than superior fourth and sixth grade boys (.05 level). Comparisons between average boys and girls and slow boys and girls in the ability grouped district, however, indicated no significant sex differences although there was a trend towards more overachievers and fewer underachievers among the girls.

In the random grouped sample, sex differences between pupils of similar ability levels were somewhat more pronounced. For superior students, the difference was significant at the .05 level. A highly significant difference occurred between the sexes for average pupils in District R. Differences in the proportion of over- and underachievers between slow boys and girls in the control group, however, were not significant. The reader is referred to Table 5-3 for a summary of differences in the proportions of overachievers, normal achievers and underachievers among the various boys and girls subgroups in the study.

## Grade Differences

The next step in the analysis involved determining whether differences existed between the proportion of overachievers, normal achievers and underachievers in the fourth grade as contrasted with the sixth grade samples within each grouping treatment. An examination of the chi square contingency tables revealed that in every case there was a trend toward fewer overachievers and a greater proportion of underachievers from fourth to sixth grades. In District A, this tendency was significant at the .01 level for both boys and girls (chi square equaled 12.67 and 9.47 respectively). In District R, the same trend existed although it was somewhat more marked in the case of the girls (chi square for boys equaled 11.95; for girls, 22.21). This trend towards a drop in the proportion of pupils who were achieving up to or beyond expectation and increase in the number of pupils who were achieving below expectation may have been a result of a change in the pupils' value system that occurred between these two grade levels. At the fourth grade level, most pupils are still oriented strongly towards the teacher and are likely to value school achievement more highly because of the teacher's influence. By sixth grade, however, the pupil's orientation has shifted to a considerable extent to the peer group in which school achievement may be valued very little or may even be ridiculed as a sign of submission to adult authority.

## Between-treatment Differences

Although the sex differences and grade differences we have discussed provide useful insights into the nature of underachievement, the main purpose of this phase of the study was to compare the proportion of overachievers, normal achievers, and underachievers found in the two grouping treatments.

A noteworthy finding in comparing underachievers in the two systems of education is the markedly greater percentage of underachievers in the random grouped system. Of the 793 pupils in District R, 175 or 22 per cent were classified as underachievers. In contrast, only 105 of the 674 District A cases or 15.6 per cent were classified as underachievers. This difference is highly significant. As the courses of study and overall achievement levels within the two districts were very similar, the comparatively large number of underachievers in the random grouped sample may be an important factor in evaluating the relative merits of random versus ability

Table 5-3

Differences Between Boys and Girls in the Distribution of Overachievers, Normal Achievers and Underachievers

| | Favored Group[a] | Chi Square | Significance Level |
|---|---|---|---|
| DISTRICT A | | | |
| Superior Boys vs. Superior Girls | G | 7.58 | .05 |
| Average Boys vs. Average Girls | G | 3.87 | NS |
| Slow Boys vs. Slow Girls | G | 3.56 | NS |
| Boys vs. Girls (overall) (4th Gr.) | G | 8.45 | .05 |
| Boys vs. Girls (overall) (6th Gr.) | G | 16.07 | .001 |
| DISTRICT R | | | |
| Superior Boys vs. Superior Girls | G | 8.63 | .05 |
| Average Boys vs. Average Girls | G | 22.23 | .01 |
| Slow Boys vs. Slow Girls | G | 3.00 | NS |
| Boys vs. Girls (overall) (4th Gr.) | G | 19.04 | .001 |
| Boys vs. Girls (overall) (6th Gr.) | G | 17.36 | .001 |

[a] "G" indicates girls had a more favorable distribution of overachievers, normal achievers and underachievers.

grouping. Comparisons of the number of overachievers shows even greater differences. The ability grouped sample had 29.1 per cent overachievers as compared with 18.9 per cent in the random grouped sample. This difference is highly significant beyond the .001 level.

The proportions of overachievers and underachievers among ability grouped and random grouped subsamples of comparable ability also showed some interesting differences (see Table 5-4). Examination of comparable subsamples indicated that the percentages of boys and girls in the districts were generally similar. Therefore, for this phase of the analysis, the sexes have been combined within each subsample. Fourth grade and sixth grade groups have been analyzed separately so that the results for one grade could be checked against the results for the other grade, and so that trends could be studied. To compare the proportions of overachievers and underachievers in comparable ability level subgroups of the two districts, chi square contingency tables were set up in which superior, average and slow students in the ability grouped sample were compared with the same subgroups in the random grouped sample.

The comparison between superior pupils showed that the ability grouped sample had a greater proportion of overachievers and a smaller proportion of normal achievers and underachievers than the random grouped sample. This difference was significant at the .05 level for the fourth grade sample. In this sample, 47.2 per cent of the ability grouped pupils were classified as overachievers as compared with 31.6 per cent of random grouped pupils. Greater differences in the same direction were found in comparing the superior sixth grade pupils in the ability grouped and random grouped samples with 41.3 per cent of the ability grouped superior pupils classified as overachievers as contrasted with 13 per cent of the superior pupils in random grouped classes. Differences between the two systems in the proportions of sixth grade superior students classified as underachievers were smaller with 16.7 per cent of the experimental group and 20 per cent of the control group so classified.

The proportions of overachievers, normal achievers and underachievers among average pupils in the two districts were not significantly different for either the fourth or sixth grade samples. A larger proportion of underachievers was found among the average pupils in random grouped classrooms but this difference was not large enough to establish a significant overall trend.

A comparison of slow fourth grade pupils in ability grouped and random grouped classes also showed no significant difference in the proportion of overachievers, normal achievers, and underachievers. These results were supported by comparison of slow sixth grade pupils in the two samples which also failed to reveal significant differences.

## Comparisons Between Normal Achievers and Underachievers

In addition to comparing the proportions of overachievers and underachievers in each district, data were collected on a number of dependent variables hypothesized to be related to underachievement. Four major subgroups of pupils from Sample VI were involved in these comparisons: underachievers in ability grouped classes, underachievers in random grouped classes; normal achievers in ability grouped classes and normal achievers in random grouped classes. Each of these four groups, in turn, was divided into six subgroups on the basis of ability level and sex.

In analyzing scores on the dependent variables, overall differences between underachievers and normal achievers in each district and between underachievers in the two

Table 5-4

Differences Between District A and District R Pupils in the Distribution of Overachievers, Normal Achievers and Underachievers

| | Favored Group | Chi Square | Significance Level |
|---|---|---|---|
| 4th Gr. Superior Pupils | A | 8.43 | .05 |
| 6th Gr. Superior Pupils | A | 41.14 | .001 |
| 4th Gr. Average Pupils | A | 3.93 | NS |
| 6th Gr. Average Pupils | A | 1.76 | NS |
| 4th Gr. Slow Pupils | A | 3.94 | NS |
| 6th Gr. Slow Pupils | A | 4.68 | NS |

districts and normal achievers in the two districts were made. Comparisons by ability level were also made; that is, superior underachievers in the two districts were compared with each other and superior underachievers in each district were also compared with superior normal achievers. Sex difference comparisons were also made between pupils in the same treatment, achiever classification and ability classification. In carrying out the analysis, sex difference comparisons were generally made first and if no sex differences or only a few inconsistent sex differences emerged, the sexes were combined in order to provide larger samples for the remainder of the analysis.

## Study Methods

Normal achieving and underachieving pupils in both districts were administered the California Study Methods Survey (CSMS) during the seventh grade. The total score on this measure was employed in the statistical analysis. The first comparisons made were to determine whether sex differences were present between comparable groups within each district. No significant differences occurred between boys and girls in comparable groups in the District A sample. In District R, there were no significant sex differences for the underachieving pupils but among normal achieving pupils, both superior and slow girls obtained more favorable scores than comparable boys. As only two of the twelve between-sex comparisons yielded significant differences, boys and girls were combined for the remainder of the analysis of the CSMS data. The results of these comparisons may be found in Figure 5-1. This figure lists the 12 subgroups compared and the lines between various subgroups indicate whether differences between the groups they connect were statistically significant. For example, at the top of the figure, the line between District A underachievers and District R underachievers is marked not significant which indicates that overall comparisons between the two underachiever groups did not yield significant differences on the CSMS. The horizontal lines in Figure 5-1 represent differences between the two treatments, the long vertical lines represent differences between comparable groups of normal achievers and underachievers within each treatment and the short vertical lines represent differences between pupils at the three ability levels within each of the achievement catagories.

It may be seen that of the eight between-treatment comparisons on the California Study Methods Survey, only one, the overall difference between normal achievers in the two districts reached statistical significance. These results suggest that pupils of comparable ability level and of comparable achievement catagory, were not significantly different in their study methods.

Of the total of 12 comparisons of pupils at different ability levels but in the same treatment and achiever catagories, four reached statistical significance. All of the significant differences obtained were between pupils of different ability levels in District R and generally reflected better study habits for superior pupils as compared with average and slow pupils in both the underachiever and normal achiever subsamples. No significant differences in study methods emerged in comparisons between District A pupils of different ability levels.

Perhaps the most striking results obtained on this measure was the lack of significant differences between normal achievers and underachievers. A total of eight comparisons were made between comparable ability classifications of underachievers and normal achievers in each district. None of these eight differences reached statistical significance. It had been hypothesized that normal achievers would generally obtain more favorable scores on the CSMS than underachievers. In view of the consistent lack of significant differences in both districts, this hypothesis must be

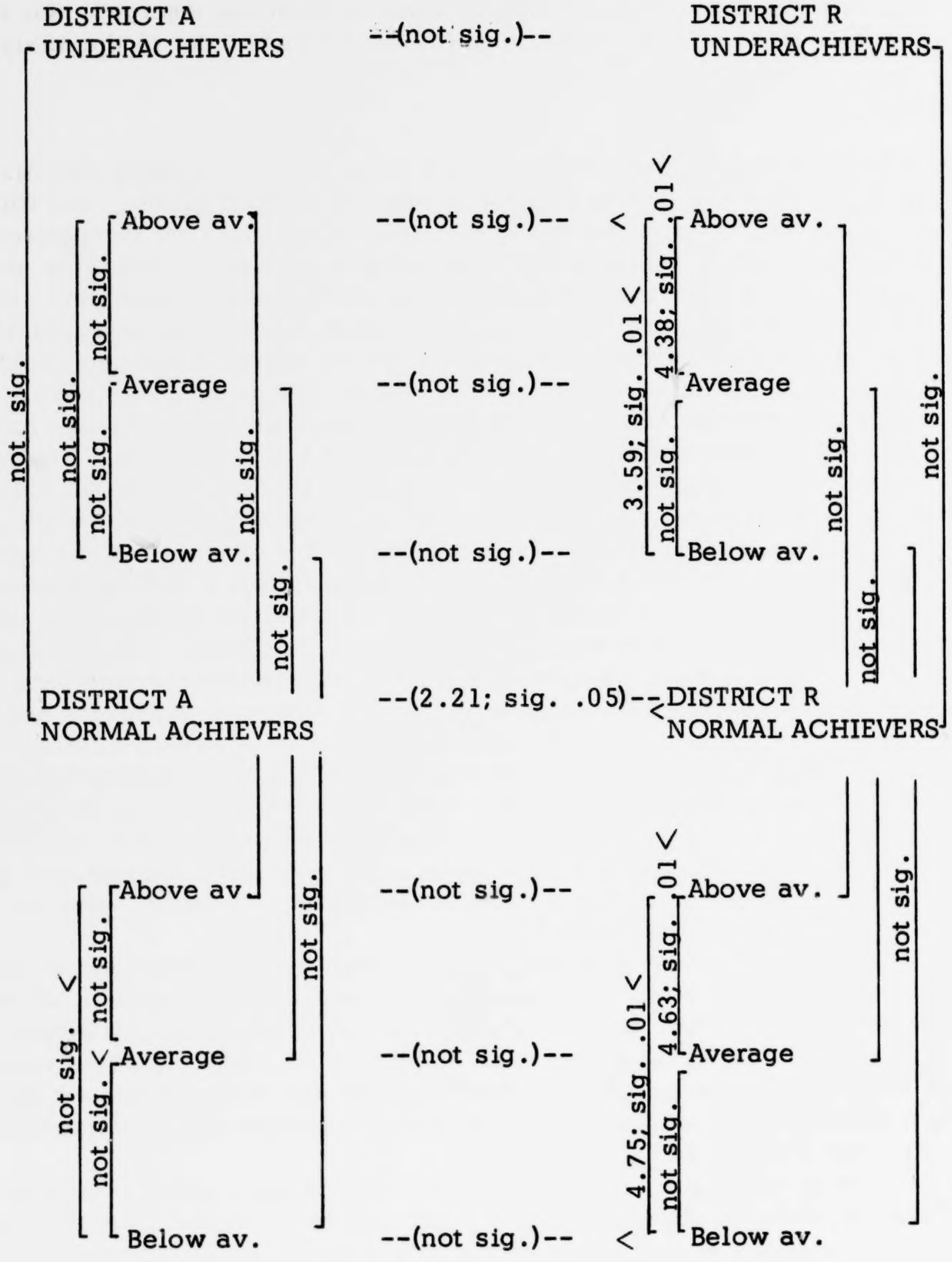

DISTRICT A UNDERACHIEVERS
--(not sig.)--
DISTRICT R UNDERACHIEVERS
Above av.
--(not sig.)--
<
.01 <
Above av.
Average
--(not sig.)--
4.38; sig.
3.59; sig. .01 <
Average
Below av.
--(not sig.)--
Below av.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
DISTRICT A NORMAL ACHIEVERS
--(2.21; sig. .05)--
<
DISTRICT R NORMAL ACHIEVERS
Above av
--(not sig.)--
.01 <
Above av.
<
v
Average
--(not sig.)--
4.63; sig.
4.75; sig. .01 <
Average
Below av.
--(not sig.)--
<
Below av.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.

Figure 5-1. Comparisons and results of the California Study Methods Survey

rejected and it must be concluded that study methods as measured by the CSMS was not a significant factor in underachievement for the samples employed in this research.

## Emotional Disturbance

It was hypothesized that higher levels of emotional disturbance would generally be found among underachievers than among normal achievers. To test this hypothesis, the Thinking About Yourself questionnaire developed by the California State Department of Education was employed. As no significant sex differences were found between comparable boys' and girls' groups in either District A or District R, the sexes were combined for this analysis. Comparisons between the mean scores of pupils in the two different grouping treatments, at different ability levels, and of different achievement classifications, are summarized in Figure 5-2, following the same pattern as employed in Figure 5-1. High scores on this measure indicate a greater degree of emotional disturbance and are considered undesirable.

Between-treatment comparisons showed two significant differences among the eight comparisons made. Superior District A underachievers had significantly higher emotional disturbance scores than comparable District R underachievers. Slow District A normal achievers had higher emotional disturbance scores than did slow District R normal achievers. The non-significant between-treatment differences also reflected a general trend for District A groups to show a greater amount of emotional disturbance than District R groups.

Comparisons between underachievers at different ability levels revealed one significant difference indicating higher emotional disturbance scores for superior District A underachievers than for slow underachievers in the same district. Comparisons between normal achievers at different ability levels in District A show two of the three differences to be statistically significant, all reflecting higher emotional disturbance for groups of lower ability. Emotional disturbance on this measure is indicated by the discrepancy between "self" and "wanted self." Thus, the greater number of level differences for normal achievers may reflect a tendency for lower ability pupils who are achieving normally to be somewhat disturbed because of their inability to compete with more able students. On the other hand the absence of such disturbance among underachievers may indicate that these pupils have developed a value system in which achievement is not an important aspect of self-concept. No significant level differences were found in District R among either normal achievers or underachievers. This may reflect less worry among District R pupils concerning their achievement or perhaps indicates that District R pupils have a less realistic perception of their performance.

In comparing similar groups of normal achievers and underachievers in each district, only one significant difference was found. This difference shows somewhat higher levels of disturbance among below average normal achievers in District A as compared with below average underachievers in the same district. This difference may reflect the value system difference previously mentioned or may indicate that the below average normal achievers in District A are, in some way, a biased sample. When it is considered that this subsample is involved in four of the seven significant differences shown in Figure 5-2 and in view of the fairly small size of this sample, it may well be that these differences reflect a sampling bias. An alternate interpretation would be that ability grouping is especially related to high levels of emotional disturbance among slow pupils who are realizing a normal achievement potential.

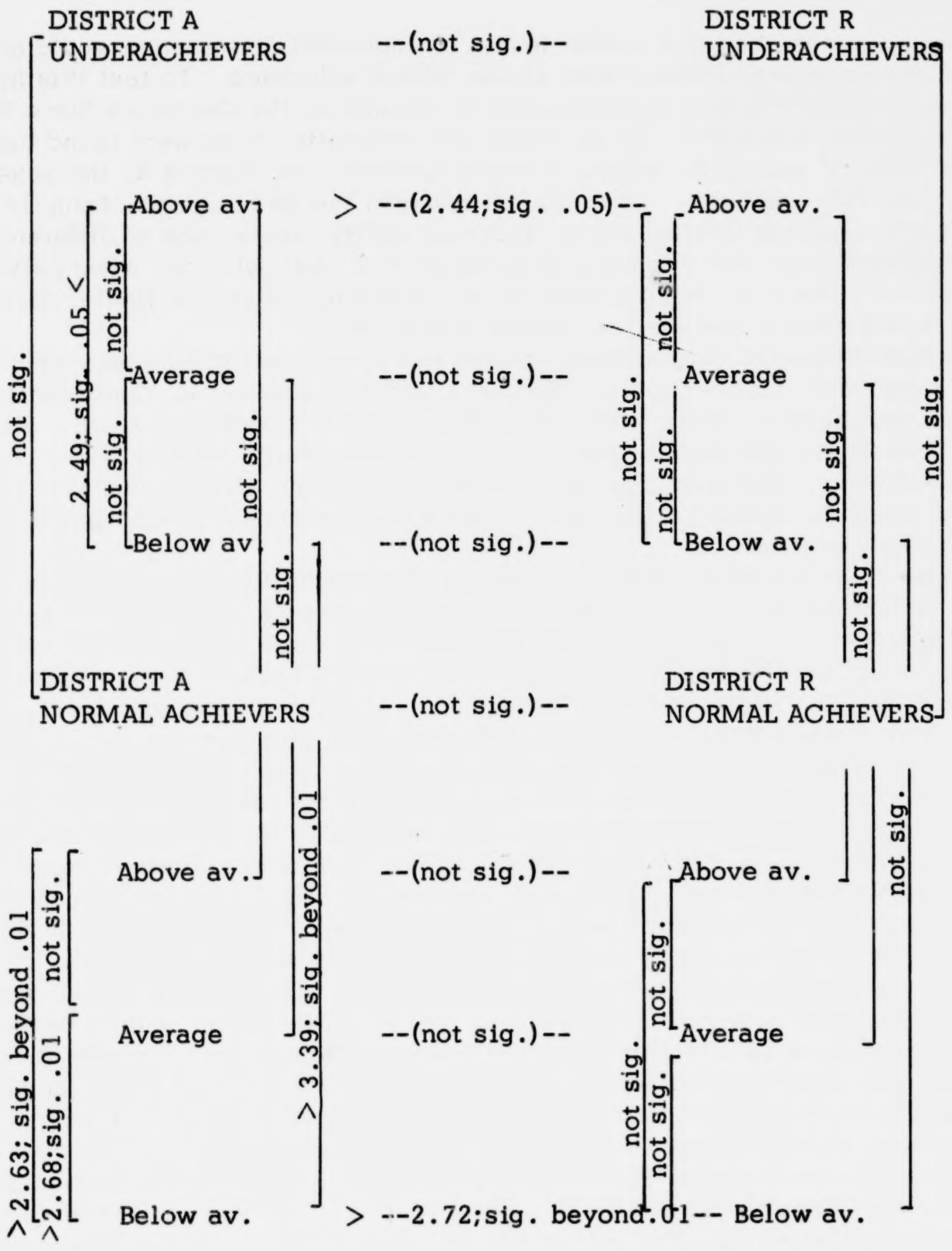

DISTRICT A
UNDERACHIEVERS
--(not sig.)--
DISTRICT R
UNDERACHIEVERS
Above av.
> --(2.44;sig. .05)--
Above av.
not sig.
2.49; sig. .05 <
not sig.
Average
--(not sig.)--
Average
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
Below av
--(not sig.)--
Below av.
not sig.
not sig.
DISTRICT A
NORMAL ACHIEVERS
--(not sig.)--
DISTRICT R
NORMAL ACHIEVERS
Above av.
--(not sig.)--
Above av.
> 2.63; sig. beyond .01
not sig.
> 2.68;sig. .01
> 3.39; sig. beyond .01
not sig.
Average
--(not sig.)--
Average
not sig.
not sig.
not sig.
Below av.
> --2.72;sig. beyond .01-- Below av.

Figure 5-2. Comparisons and results of the "Thinking About Yourself" test

## Achievement Motivation

Achievement motivation was measured by analyzing the subject's responses to four Thematic Apperception Test cards and scoring these for achievement motivation using Atkinson's (1958) system. No sex differences were found between comparable groups so boys and girls were combined for the remainder of the analysis. The results of the analysis of the achievement motivation scores are shown in Figure 5-3. It was hypothesized that underachievers would obtain lower scores on this variable than would normal achievers. Perhaps the most noteworthy finding was the general lack of significant differences on this variable. None of the between-treatment comparisons yielded statistically significant differences. The 12 comparisons between pupils in the same treatment and achievement classification but at different ability levels yielded two significant differences and three differences approaching significance, all favoring the group of higher ability. None of the eight comparisons between underachievers and normal achievers in the two districts yielded significant differences. Thus, the achievement motivation score appeared to be more successful in differentiating between pupils at different ability levels than between underachievers and normal achievers. It must be concluded that achievement motivation as measured in the Utah study was not related to underachievement. These results appear to cast doubt on the validity of the Atkinson scoring system for use with pupils at the junior high school level.

## Socioeconomic Status

Each student in this phase of the study was asked questions concerning the occupation and education of the chief wage earner of his family. The pupils' responses were given a rating from 1 to 7 on the Hollingshead Occupational Scale and Education Scale. From these weighted scales, an index of social position score (Hollingshead and Redlich, 1958) was used to arrive at a social class status. Pupils were placed in five social class catagories based on the combined weighted occupational and education scale. Differences between comparable boys and girls were then analyzed. As only one significant difference in the 12 such comparisons emerged, underachieving District A average boys were significantly higher at the .05 level than underachieving District A girls. Thus, the sexes were combined for the remainder of the analysis.

The reader will recall that the Hollingshead Occupational Scale only was employed in checking the social class comparability of other samples in the Utah study. Thus, as this phase of the study employed random samples of pupils selected from Sample VI and used a closely related measure of socioeconomic status, no differences in socioeconomic status would be expected. The expected result was borne out as no significant differences between districts were found on socioeconomic status.

Within-treatment comparisons between pupils of different ability levels yielded a number of significant differences. In District A, superior underachievers had significantly higher socioeconomic status scores than slow underachievers. This was the only significant difference in the six level comparisons for District A. In District R, superior underachievers were again significantly higher on this variable than slow underachievers at the .05 level. Comparisons between District R normal achievers showed all three of the within-treatment comparisons to be statistically significant, the superior pupils having the highest score followed in order by the average and slow pupils.

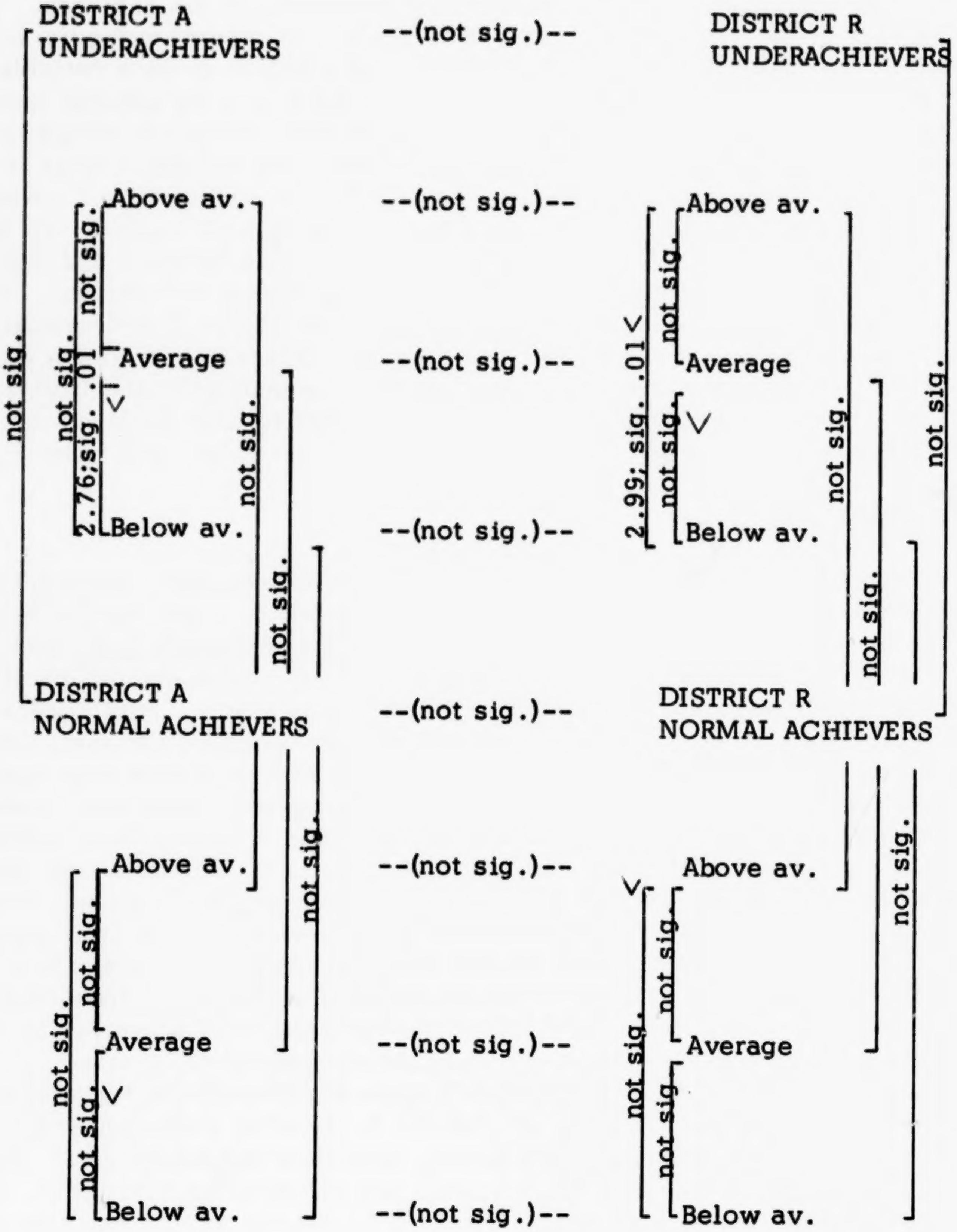

DISTRICT A UNDERACHIEVERS
--(not sig.)--
DISTRICT R UNDERACHIEVERS
Above av.
--(not sig.)--
Above av.
Average
--(not sig.)--
Average
Below av.
--(not sig.)--
Below av.
DISTRICT A NORMAL ACHIEVERS
--(not sig.)--
DISTRICT R NORMAL ACHIEVERS
Above av.
--(not sig.)--
Above av.
Average
--(not sig.)--
Average
Below av.
--(not sig.)--
Below av.
not sig.
not sig.
not sig.
2.76;sig. .01
not sig.
not sig.
2.99; sig. .01
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
not sig.
<
<
>
>
>
>

Figure 5-3. Comparisons and results of the TAT (Achievement Motives) test

Perhaps the most striking results in regard to socioeconomic status was that none of the comparisons between normal achievers and underachievers in either grouping treatment were statistically significant. This finding is generally not in agreement with the previous research (Goldberg et.al., 1959; Gibboney, 1959). The results of the Utah study may reflect the unique characteristics of the sample involved. Subjective observation indicates that even among lower class parents many children receive considerable encouragement to remain in school and achieve. This subjective observation is borne out to some extent by the fact that a very high percentage of Utah children complete high school. It may well be, however, that some of the previous research involving samples from large cities has also presented a distorted picture of the importance of socioeconomic status to achievement.

## Parents' Attitude Towards School

An estimate of the parents' attitude toward school was obtained through structured personal interviews of normal achievers and underachievers conducted by one of the investigators. The pupils' responses to the four questions concerned with parental attitude were scored and combined scores were compared for the normal achiever and underachiever subgroups in each district. The reader should note that these scores are based on relatively little evidence and that no follow-up in depth of the pupils' responses was attempted in the interview. As the questions were quite direct, however, it may be expected that the results obtained were reliable although depth interviewing may have provided greater insight into the variables being studied. Mean scores for various subgroups were compared using t-tests. No significant differences emerged in any of the comparisons of parental attitude. A total of 24 comparisons were made on this variable between normal achievers and underachievers, between the treatments, and between pupils of different ability level within each treatment. Perhaps the lack of differences between the attitudes of parents of normal achievers and underachievers indicates that for this sample, parental attitude was not a significant variable.

These results, however, indicate a somewhat more favorable mean parental attitude than would be expected in a heterogeneous population. Out of a possible score of 4.00, parental attitudes for District A underachievers yielded a mean score of 3.41 as compared with 3.44 for normal achievers. In District R the means were 3.73 and 3.70 respectively. These data, of course, represent the pupil's perception of his parents' attitudes, but in general suggest that consistently favorable attitudes (at least as reported by the pupil) are present. Because of the restricted variability of the responses, it must be concluded that these data do not provide a generalizable test of the hypothesis that socioeconomic status and parental attitudes are related to underachievement.

## Health

To determine whether health played a significant role in underachievement, pupils were given thorough physical checkups by a practicing pediatrician. The selection of possible health related causes of underachievement was made by the investigators in cooperation with the supervisor of the health department serving the respective districts, the physician in charge of the health department and by the pediatrician who was employed to carry out the medical checkups. Prior to the physical examinations, case histories were obtained on each pupil by nurses of the health department. Major items reported were: normal or abnormal delivery and development; communicable disease history and complications; health habits, including rest, food, and sleeping practices; observed listlessness or lack of vitality; allergies; operations; speech, hearing and vision; chronic diseases of all types; treatment of ailments; traumatic emotional experiences; sexual development; epilepsy and any other information which the parents wished to divulge.

The health examination itself covered the items shown in Table 5-5 plus written observations by the doctor whenever abnormalities were present. The nurses again assisted with these examinations, and each report was dated and signed by the physician

In the statistical analysis of the data, a dichotomy was established with those students who had three or more health problems were classified as "health problems cases" while those with two or fewer health problems were classified as "no health problem." Roughly 74 per cent of the students were classed as normal, having either no health problems whatsoever, or not more than two health problems.

This artificial dichotomy was employed as a means of testing the statistical significance of differences using chi square contingency tables. The chi square analysis yielded no significant difference between any of the subgroups compared on the health data. There were no differences between comparable boys and girls, no differences between pupils at different ability levels, and no differences between underachievers or normal achievers in either district.

## SUMMARY AND CONCLUSIONS

### First Phase

On the basis of the findings, the following conclusions can be drawn from the first phase of the study:

1. There was a significant tendency for girls to more frequently be overachievers and less frequently be underachievers than boys. This trend was present at both grade levels and in both treatments.

2. A comparison of fourth and sixth grade pupils in each of the grouping treatments indicated a significant trend for both boys and girls to achieve less at the sixth grade level relative to their potential. In all cases, the sixth grade samples had a greater number of underachievers and a smaller number of overachievers than comparable fourth grade samples. This drop is considerably greater for girls in District R than for girls in District A which might suggest that ability grouping tends to lessen the trend toward lower achievement between the fourth and sixth grades. This implication is not borne out by the results of the boys, however, whose drop in relative achievement is substantially the same in both the random grouped and the ability grouped samples.

3. Overall comparisons between Districts A and R indicated that ability grouped pupils were more often classified as overachievers and less often classified as underachievers than were random grouped pupils.

4. In comparing pupils at the three ability levels in the two districts, the ability grouped samples were generally found to have a significantly more favorable distribution of overachievers and underachievers at both grade levels. The differences between sixth grade superior pupils in the two districts were particularly marked with

Table 5-5

Tabulation of Health Data

| | Number of Cases | | | | | |
|---|---|---|---|---|---|---|
| | District A Achievers | | District R Achievers | | | |
| | Under | Normal | Under | Normal | Totals | Per Cent |
| Urine | 13 | 11 | 9 | 3 | 36 | 20.0 |
| Hemoglobin | | | | | -- | ---- |
| Blood Pressure | | | | | -- | ---- |
| Eyesight | 11 | 10 | 13 | 5 | 39 | 22.0 |
| Height | 18 | 12 | 13 | 13 | 56 | 31.0 |
| Weight | 23 | 14 | 16 | 17 | 70 | 39.0 |
| Ears | 1 | 2 | | | 3 | 2.0 |
| Nose | 1 | 1 | | | 2 | 1.0 |
| Throat | 3 | 1 | | | 4 | 2.0 |
| Lungs | | | | | -- | ---- |
| Heart | | 1 | | 1 | 2 | 1.0 |
| Abdomen | | | | | -- | ---- |
| Hernia | | 1 | | | 1 | .5 |
| Extremities | 3 | 1 | 3 | | 7 | 4.0 |
| Case history complications | 13 | 10 | 7 | 5 | 35 | 20.0 |
| Pubertal changes | 5 | 3 | | 1 | 9 | 5.0 |
| Orthodontic (braces) | | | 2 | | 2 | 1.0 |
| Acne, eczema or other scars, blemishes, or obvious defects | 4 | 9 | 5 | 9 | 27 | 15.0 |
| Total Sample | 87 | | 91 | | 178 | |

District A having three times as many overachievers as District R and a smaller percentage of underachievers. These results seem to support the often stated contention that ability grouping presents a greater challenge to superior students.

5. No significant differences were found between the proportions of overachievers, underachievers, and normal achievers among pupils of average ability in the two districts. These results support the working hypothesis and suggest, as might be expected, that the two systems of grouping have little effect on the achievement or motivation of average pupils. In random grouping systems the curriculum is generally aimed at the average pupil. Thus, one of the oft-stated advantages of the ability grouping, that it fits the curriculum more accurately to the needs of the pupil, does not apply to the average groups.

6. Comparisons of the proportions of overachievers, underachievers and normal achievers among slow pupils in the two systems revealed no significant differences. For slow pupils, the ability grouping system is generally thought to have the advantage of presenting the student with more suitable curricular materials while the random grouping system might be expected to produce greater achievement because of the pressure put on the slow pupil to keep up with the rest of the class. If these factors are operating, it must be concluded that neither is of sufficient importance to lead to significantly different achievement patterns for the slow pupil.

The reader is reminded once again that the above data are based upon achievement data collected during the pilot study year of the Utah project and should therefore be considered tentative until supported by additional research evidence.

## Second Phase

The findings of the second phase of the study support the following conclusions:

1. When pupils of similar ability were compared, neither grouping treatment was consistently related to better study methods.

2. Study methods were significantly related to ability level in District R but not in District A.

3. There were no significant differences between the study methods scores of normal achievers and underachievers of comparable ability. These findings lead to the conclusion that study methods as measured by the CSMS is not a significant factor in underachievement.

4. There was a consistent tendency for District A pupils to have higher emotional disturbance scores than District R pupils on the Thinking About Yourself questionnaire. This trend was present among both normal achievers and underachievers. These data suggest that ability grouping leads to greater discrepancies between "self" and "wanted self" than random grouping. It may be that the ability grouping situation forces the pupil to made a more realistic self appraisal of his school performance that in turn tends to lower his satisfaction with his performance.

5. Within-treatment comparisons showed a consistent tendency for pupils of lower ability classification in District A to have less favorable emotional disturbance scores. No such pattern emerged in the District R data. This finding appears to support the aforementioned conclusion that ability grouping makes the pupil of lower ability more aware and less acceptant of his achievement status.

6. The general lack of significant differences between the emotional disturbance scores of comparable groups of normal achievers and underachievers leads to the conclusion that this variable, as measured by the Thinking About Yourself questionnaire,

is not related to underachievement.

7. Achievement motivation was found to be related to ability level to some degree but scores on this variable did not differentiate between normal achievers and underachievers in the Utah study. These results may indicate that Atkinson's (1958) scoring system, as employed in the Utah study, is not a valid measure of achievement motivation.

8. Socioeconomic status was not related to underachievement in the Utah study. This finding may reflect differences between this and other studies in the nature of the populations sampled.

9. Parental attitudes towards school, as reported by pupils interviewed, were not significantly different for normal achievers and underachievers. Reported attitudes were high and showed little variability. These results may indicate that, in the population sampled, parental attitudes towards school are so consistently favorable as to be unimportant as a factor in underachievement. It may be, however, that the direct structured interview technique was not appropriate and yielded invalid data.

10. Incidence of health problems was unrelated to grouping treatment, sex, ability level or underachievement. It may be concluded that the physical and health variables measured in the Utah study are not a factor in underachievement.

## REFERENCES

Armstrong, Marion E. A comparison of the interests and social adjustments of underachievers and normalachievers at the secondary school level. Doctoral dissertation, University of Connecticut: University Microfilms, Ann Arbor, Michigan, 1955.

Atkinson, J. W. *Motives in fantasy, action, and society.* New York: D. Van Nostrand Company, Inc., 1958.

Bedoian, V. H. Social acceptability and social rejection of the underage, at age, and overage pupils in the sixth grade. *J. educ. Res.*, 1954, 47, 513-20.

Bower, E. M., Tashnovian, P. J., & Larson, C. A. A process for early identification of emotionally disturbed children. *Bull. Calif. St. Dept. of Educ.*, 1958, 27, 18.

Brookover, W. B., Velinsky, Ann, & Thomas, S. *The relationship of self image to achievement in junior high school students*. Unpublished manuscript, Michigan State University, East Lansing, Michigan, 1961.

Burdick, H. A. Interrelationships of three measures of motivation. *Psychol. Rev.*, 1961, 68, 225.

Fliegler, L. A. *Psychology*. (3rd ed.) New York: World Book Company, 1957.

Frandsen, A. N., & Higginson, H. B. The Stanford-Binet and the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.*, 1951, 15, 236-38.

Gibboney, R. A. Socioeconomic status and achievement in social studies. *Elem. Sch. J.*, 1959, 59, 340-46.

Goldberg, Miriam L., *et. al*. A three year experimental program at DeWitt Clinton High School to help bright underachievers. *High Points*, 1959, 41, 1.

Goodlad, J. I. Pressures can be blocks to learning, don't push me. *Bull. Assoc. for Childh. Educ. International.*, 1960, 23-8.

Hinkelman, E. A. Relationship of intelligence to elementary school achievement. *Educ. Admin. Superv.*, 1955, 41, 176-79.

Hollingshead, A. B., & Redlich, R. C. *Social class and mental illness*. New York: John Wiley and Sons, Inc., 1958.

Jensen, V. H. Influence of personality traits on academic success. *Personnel Guid. J.*, 1958, 36, 497-500.

John, L. The relationship of the achievement motives as projected into Thematic Apperception stories to children's school achievement. Unpublished doctoral dissertation, Utah State University, Logan, 1960.

Klausmeier, H. J. Physical, behavioral, and other characteristics of high- and lower-achieving children in favored environments. *J. educ. Res.*, 1958, 51, 573-81.

MacLean, M. S. What and why of underachievement. *H. Sch. J.*, 1958, 42, 69-72.

Martinson, R. A. Study project on programs for gifted pupils. Unpublished manuscript, California State Department of Education, Sacramento 14, California, 1960.

Muggenthaler, E. The sociological status of slow-learning pupils in an elementary grade. *Prax, Kinder-psychologie Kinder-psychiatrie*, 1955, 4, 226-30.

Rosenfeld, H., & Zander, A. Teachers' influence on student aspirations. *J. educ. Psychol.*, 1961, 52, 1-11.

Russell, D. W. A plea to beam in the underachiever. *H. Sch. J.*, 1958, 42, 66-8.

Sears, P. S. Levels of aspiration in academically successful and unsuccessful children. *J. abnorm. Soc. Psychol.* , 1940, 35, 498-536.

Stivers, E. H. Motivation for college in high school girls. *Sch. R.*, 1959, 67, 320-34.

Stone, D. R., & Ganung, G. R. A study of scholastic achievement related to personality as measured by the Minnesota Multiphasic Personality Inventory. *J. educ. Res.*, 1956, 50, 155-56.

Ullmann, C. A. Identification of maladjusted school children. *Public Health Monograph*, 1952, No. 7.

Williams, Meta F. Acceptance and performance among elementary school children. Ed. Digest, 1959, 24, 41-43.

Winkler, J. A., & MacNutt, T. Problems involved in underachievement at the fourth grade level. in Research Relating to Children. Bulletin No. 11, U.S. Department of Health, Education, and Welfare. Washington 25, D.C.: U.S. Government Printing Office, 1960.

## Chapter 6

# STUDY HABITS

### INTRODUCTION

Although no previous research was found that was concerned directly with the effects of ability grouping and random grouping upon study habits, this area seemed a worthwhile one to explore. Ability grouping--coupled with acceleration as practiced in District A--would seem to place pressure upon pupils to achieve and in so doing, might be expected to encourage better study habits. On the other hand, an enrichment program such as that employed in District R would provide bright pupils with the time and materials for independent study which could develop better study habits than the pressures of the accelerated program.

### WORKING HYPOTHESES

Many educators are convinced that ability grouping tends to develop better study habits. This belief is particularly strong with regard to superior pupils and is based on the idea that the superior pupil in a random grouped class is not sufficiently challenged and is capable of doing the assigned work with very little real effort. Thus, he tends to develop poor study habits and attitudes. Conversely, in an ability grouped classroom the bright student is forced to compete with other students of equal ability in a more challenging program, requiring him to develop better study habits in order to maintain his position in the class. Thus, our working hypothesis was that superior pupils in ability grouped classrooms would obtain significantly higher scores on the study habits measures than would superior pupils in random grouped classrooms.

There is little reason to expect that ability grouping would tend to develop better study habits among average or slow pupils. In fact, it might be argued that the average pupil would be motivated to study harder in the random grouped class where he is challenged by superior pupils. With regard to the slow pupil the same argument can be made but a counter argument is also possible to the effect that the slow pupil in the random grouped class is faced with competition he cannot hope to meet and may thus react with some form of rationalization which would tend to reduce his efforts rather than increase them.

In view of the lack of evidence in this area and the absense of factors in the grouping treatments that would suggest a clear cut superiority for either treatment, it was hypothesized that there would be no significant differences between the scores of average and slow pupils on the study habits measures.

## PROCEDURE

### Sample

Considerable evidence concerning study habits was collected over the four years of the Utah study. Pupils in Sample IV, VI, VIII, and IX were administered study habits measures at the grade levels indicated in Table 6-1.

### Measures

#### California Study Methods Survey (CSMS)

This questionnaire contains 150 items and yields four subscores and a total score. The first subscore on the CSMS attempts to measure Attitudes Toward School. Sixty items of the scale deal with this area and attempt to measure ". . . the level of the student's morale and his feelings of harmony with the school community" (Carter, 1958, p. 3). The second subscore, Mechanics of Study also contains 60 items and attempts to get at the effectiveness of the student's study techniques and his methods of performing the mechanical aspects of study such as use of outlines, notetaking, reviewing for tests, and other specific study techniques. The third sub-score, Planning and System contains 30 items and attempts to measure the extent to which the student budgets his study time and the degree of planning, care, and deliberation he applies to his work. A total score based on the entire 150 items may also be calculated and provides an overall estimate of the students' study methods.

In addition to these four scores, a verification key is also available. The test manual suggests a critical score of 17 and indicates that verification scores below this point indicate that the student is either attempting to fake an unduly favorable score, or is careless in his marking of the test, or has omitted too many items. In this research, pupils below the suggested critical score were eliminated prior to carrying out the analysis.

#### Survey of Study Habits and Attitudes

Unlike the CSMS, the Brown Holtzman Survey yields a total score only. It contains 75 items, about half dealing with study methods and procedures and half dealing with study attitudes. The test-retest reliability of the total score is reported as .83 for boys and .80 for girls (Brown and Holtzman, 1956) which is comparable to the total score reliability of .86 for CSMS (Carter, 1958). The principal reason for using the SSHA with Samples VIII and IX was the shorter testing time, (20 minutes versus 40 minutes of CSMS) which is an important consideration when subjects are to be administered an extensive battery.

## RESULTS

### Sample IV Data

#### Introduction

Because of the lack of suitable study habits measures that can be administered at the elementary school level, Sample IV pupils were not administered a study habits

Table 6-1

Samples Employed in the Study Habits Phase of the Utah Study

| Sample | N | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|
| IV | 727 | CSMS | | | | | |
| VI | 673 | | | CSMS | | | |
| VIII | 428 | | | SSHA | | SSHA | |
| IX | 214 | | | SSHA | SSHA | | SSHA |

test until the final year of the study when they were in grade seven. Near the end of grade seven, these pupils were administered the California Study Methods Survey (CSMS). This measure contains 150 items and yields four scores plus a total score. During three years of the four devoted to this research, Sample IV were pupils at the elementary school level. Thus, the scores of these pupils obtained in grade seven provide the best evidence available in this project on the relative effectiveness of random grouping with enrichment vs. ability grouping with acceleration in developing study attitudes and skills at the elementary level.

## Limitations in the Sample IV Data

As no suitable study methods measure was not available for administration to Sample IV at the beginning of the study, one cannot be certain that initial study habits differences had not already developed between pupils in the two districts prior to the beginning of the research project. In view of the fact that the elementary schools from which these samples were drawn were carefully matched between districts, and the fact that both districts had employed random grouping methods in self contained classrooms prior to the research and were found to be comparable in a great many respects, it seems reasonably safe to assume that the groups were initially comparable in terms of study habits. In any event it would be surprising to find significant study habits differences developing in two such similar school systems by the beginning of grade four. In our treatment of the study habits data for some of the other samples, however, such as Sample IX, where an initial study habits test was administered, it will be seen that some initial differences favoring District R were present. The secondary schools, being much larger than the elementary schools in the study, were more difficult to match so that it is doubtful that initial differences comparable to those found in Sample IX were present at the fourth grade level in Sample IV. The possibility, however, does exist that the Sample IV subjects in Districts A and R were initially different in study habits at the fourth grade level and this possibility should be kept in mind in interpreting the results.

## Results

Complete data necessary to analyze the results obtained from the seventh grade administration of the California Study Methods Survey to Sample IV pupils were available on 727 cases after those subjects with verification scores below 17 had been eliminated. These pupils were divided into six groups representing the three ability levels in each district. The numbers of cases in each group may be found in Table 6-2.

In comparing Sample IV subjects on the first scale of the CSMS, Attitudes Toward School, it will be noted that District R pupils earned higher scores than comparable District A pupils at all three ability levels. The difference between mean scores of superior pupils in the two districts was significant at the .01 level in favor of District R. District R average pupils were also significantly higher at the .01 level. The differences between slow pupils in the two districts on this scale were not statistically significant.

The second subtest of the CSMS Mechanics of Study also showed differences generally favoring the District R pupils (see Table 6-2). On this subtest District R pupils were significantly higher than comparable District A pupils at all three ability levels.

Table 6-2

Mean Scores of Sample IV Subgroups on the California Study Methods Survey

| Group | N | Attitudes Toward School Mean | Mechanics of Study Mean | Planning and System Mean | Total Score Mean |
|---|---|---|---|---|---|
| Ability Grouped (Dist.A) Superior | 97 | 35.59 | 40.25 | 20.51 | 96.37 |
| Random Grouped (Dist.R) Superior | 202 | 38.52[a] | 41.86[b] | 22.23[a] | 102.62[a] |
| Ability Grouped (Dist.A) Average | 148 | 31.57 | 36.32 | 19.22 | 87.10 |
| Random Grouped (Dist.R) Average | 157 | 34.43[a] | 38.30[a] | 20.57[b] | 93.29[a] |
| Ability Grouped (Dist.A) Slow | 42 | 30.33 | 32.28 | 19.14 | 81.75 |
| Random Grouped (Dist.R) Slow | 81 | 31.39 | 36.42[a] | 19.74 | 87.58[a] |
| | 727 | | | | |

[a] This indicates a difference significant at the .01 level favoring the indicated group as compared with the comparable ability level group in the other district.

[b] This indicates a difference significant at the .05 level.

Between-treatment comparisons on the third sub-scale of the CSMS, Planning and System, again showed scores generally favoring the District R groups. District R superior pupils and average pupils were both significantly higher than their District A counterparts. No significant difference was found between the mean scores of the slow groups in the two districts on this subtest.

The last analysis involving Sample IV was concerned with their performance on the total score of the California Study Methods Survey. As would be expected from the subtests already discussed, the District R pupils were again found to be superior to pupils of comparable ability level in District A. The difference between mean scores of superior pupils favored the District R subjects by a considerable margin. This difference was statistically significant beyond the .01 level. In terms of percentile norms superior District A pupils were near the top of the 55-64.9 percentile band while District R pupils are near the bottom of the 75 to 84.9 percentile band (Carter, 1958). District R pupils at the average ability level were also significantly higher than comparable pupils in District A. This difference was also beyond the .01 level and represents a difference of more than 10 percentiles based on the CSMS norms.

The difference between the slow groups also favored District R but was somewhat smaller although still statistically significant at the .05 level.

## Summary

In looking over the Sample IV results on the California Study Methods Survey one is impressed by the consistently higher scores of District R pupils. Although it is possible that some of these differences were already present at the beginning of the fourth grade it appears highly probable that at least a part of the differences that emerged are attributable to the grouping systems. District R superior pupils made significantly higher mean scores on all three of the CSMS subtest scores plus the total score as compared with District A superior pupils. Three of these differences were at the .01 level while the fourth, Mechanics of Study, was significant at the .05 level. All differences between average groups in the two districts were also statistically significant favoring District R. Among slow pupils, the Mechanics of Study mean significantly favored the District R sample at the .01 level, as did the total score. Differences between the slow groups on the other two subtests were not statistically significant.

## Discussion

It is interesting to speculate on possible reasons for the consistent superiority of the study habit scores earned by District R pupils in Sample IV. It may be that the enrichment program in effect in District R elementary schools provided the superior pupil with more time for independent study than was available in the ability grouped system with its accelerated curriculum. In the process of independent study the superior pupil in District R may have developed the better study habits and attitudes that are reflected in the CSMS scores.

This possible explanation, however, would hardly account for the differences in average pupils, because the amount of time the average pupil has to devote to enrichment materials in the heterogeneous classroom is rather limited and the materials available are usually aimed at the extreme groups with most materials for the superior pupils. These pupils, however, may have been influenced by another factor, namely the challenge presented to them by the presence of superior pupils in their class. The superior pupils may have motivated the average pupils to set higher goals, and better study habits and attitudes may have resulted. This same factor could also have operated to motivate low ability pupils although one would expect it to be less effective for this group because many of the goals and performance standards set by superior pupils in the heterogeneous class would be unattainable for pupils of low ability, and may have tended to lead to discouragement or rationalization on the part of some of these pupils.

One must conclude on the basis of the Sample IV data that ability grouping of the type used in District A does not lead to better study habits for elementary pupils at any ability level. Instead, the data indicate that it is highly probable that heterogeneous grouping with enrichment tends to develop better study habits at the elementary levels, especially among superior and average pupils.

## Sample VI Data

### Introduction

The California Study Methods Survey was administered to Sample VI pupils near the end of the ninth grade. As was the case with Sample IV pupils, no initial study habits measure was obtained. Therefore, some initial differences in study habits may have existed between the District A and District R pupils at the beginning of the study. As was the case with Sample IV, however, Sample VI pupils were initially comparable in many ways, and the curriculum and teaching methods were similar in the two districts at elementary level prior to the introduction of ability grouping in District A. Thus, there seems to be a sound basis for expecting that little or no initial difference existed between pupils of comparable ability in the two districts. During three of the four years of the study, Sample VI were pupils at the junior high school level. Thus, the scores of these pupils obtained in grade nine provide the best evidence this research has to offer on the relative effectiveness of ability grouping with acceleration in mathematics and science as contrasted with random grouping with enrichment at the junior high school level.

### Results

After eliminating incomplete cases and those who scored less than 17 on the CSMS verification score, a total of 673 complete cases were available for the analysis. An analysis of variance was carried out on each of the three subscores and the total score following the same procedure employed with Sample IV. A summary of the results obtained by comparing pupils of similar ability in District A and District R may be found in Table 6-3.

In the first subtest of the CSMS, Attitudes Toward School, we find that superior pupils and slow pupils in the two districts earned very similar scores. The mean score, however, for average pupils in District R was significantly higher at the .01 level than the mean score obtained by the District A average subsample.

On the second subtest, Mechanics of Study, we again find the only significant difference between pupils of comparable ability in the two grouping systems was at the average level where District R pupils obtained a significantly higher mean score. On the Planning and System subtest of the CSMS the same pattern continued, although

Table 6-3

Mean Scores of Sample VI Subgroups on the California Study Methods Survey

| Group | N | Attitudes Toward School Mean | Mechanics of Study Mean | Planning and System Mean | Total Score Mean |
|---|---|---|---|---|---|
| Ability Grouped (Dist.A) Superior | 69 | 35.35 | 43.09 | 20.33 | 98.91 |
| Random Grouped (Dist.R) Superior | 107 | 36.03 | 43.17 | 19.66 | 98.91 |
| Ability Grouped (Dist.A) Average | 226 | 30.21 | 38.12 | 18.13 | 86.37 |
| Random Grouped (Dist.R) Average | 165 | 33.91[a] | 40.89[a] | 19.10[b] | 93.90[a] |
| Ability Grouped (Dist.A) Slow | 29 | 29.79 | 34.48 | 18.07 | 82.34 |
| Random Grouped (Dist.R) Slow | 77 | 29.43 | 36.70 | 17.93 | 84.06 |
|  | 673 |  |  |  |  |

[a] This indicates a difference significant at the .01 level favoring the indicated group as compared with the comparable ability level in the other district.

[b] This indicates a difference significant at the .05 level.

the District R average pupils on this test were only slightly higher than the District A pupils. This difference, however, was significant at the .05 level.

Comparisons between pupils of comparable ability in Sample VI on the total score showed the two superior groups to have obtained identical mean scores. No significant difference was found between the slow pupils in the two groups on the total scores. District R average pupils were, of course, higher than the District A average pupils.

## Discussion

In looking over the differences occurring on the CSMS for Sample VI pupils in the two districts the lack of significant differences at the superior and slow level provide a sharp contrast with the Sample IV results (see Table 6-2). District R superior and slow pupils received higher scores on all three subtests and on the total score in Sample IV, while none of these differences were significant for Sample VI. For both of the samples, District R average pupils were significantly higher on all three subtests and on the total score. In studying these results several interpretations are possible. One point that seems worth mentioning is that the general lack of significant differences in Sample VI tends further to support the assumption that these pupils were not different in study habits at the beginning of the study. As the Sample IV pupils came from the same elementary schools as the Sample VI pupils, it would seem that the lack of results in Sample VI suggests that Sample IV pupils were probably similar in study habits at the beginning of the study. This would lend further strength to the tentative conclusion that the grouping system played an important role in the significant differences obtained for Sample IV pupils.

One might wonder why similar differences were not found for Sample VI. Perhaps the most likely explanation is that the Sample IV results for the most part reflect changes due to differences in the grouping systems at the elementary school level. It will be recalled that elementary schools in both districts are organized into self contained classrooms with ability grouping coupled with acceleration being used as a device for adjusting the curriculum in District A while random grouping with enrichment was the method of curriculum adjustment in District R. The Sample VI data on the other hand reflect, for the most part, the effects of the two grouping systems at the junior high school level. Prior to grade six, pupils in both districts were in random grouped self contained classrooms. During grade six, District R pupils were in ability grouped classrooms and subsequently at the junior high school level were in ability grouped classrooms in mathematics and science only. District R, it will be recalled, employed random grouping with enrichment at all grade levels. The lack of significant between-treatment differences for superior and slow pupils would suggest that at the junior high school level enrichment is less effective in developing study habits than at the elementary school level. This might be expected because the self contained classroom would seem to offer better opportunities for enrichment than the departmentalized class organization employed at the junior high school level.

Perhaps the best explanation for the consistently higher study habit scores of District R average groups in both Sample IV and Sample VI is that average pupils in heterogeneous classes are challenged to a greater extent than average pupils in classes from which the superior pupils have been removed such as are found in an ability grouping system. It is doubtful if enrichment contributed materially to the better study habits of District R average pupils at either grade level as the enrichment program tends to emphasize adjustment of the curriculum for the superior and slow pupils and as a rule little adjustment is provided for pupils of average ability.

## Sample VIII Data

### Introduction

The Brown Holtzman Survey of Study Habits and Attitudes (SSHA) was administered to Sample VIII and Sample IX pupils. As both of these samples were tested near the end of junior high school and again near the end of high school, they provide evidence primarily upon the relative effectiveness of ability grouping with acceleration versus random grouping with enrichment at the high school level. These samples were not as initially comparable between districts with respect to most of the descriptive data collected as were Samples IV and VI. Thus, the administration of the SSHA on at least two occasions was desirable as it permitted adjustment for initial differences, using covariance analysis.

As the Brown-Holtzman Survey of Study Habits and Attitudes (SSHA) is not designed for administration prior to grade nine, Sample VIII pupils were not tested on this measure during the first year of the study. They were, however, tested during their ninth grade year and again during their eleventh grade year. Complete data required for the analysis of covariance were available for 428 pupils in Sample VIII. Pupils in each district were divided into their ability level groups thus giving six groups upon which to carry out the analysis. The slow groups in both districts included 37 complete cases. These were the smallest sub-groups in the Sample VIII analysis.

### Results

In looking over the mean study habits scores obtained in the ninth grade administration of the SSHA it may be seen that the District R superior and average pupils were somewhat higher than pupils of comparable ability in District A. This difference was significant beyond the .01 level for District R average pupils but was not statistically significant for superior pupils. It will be seen in Table 6-4 that the slow pupils in the two districts had almost identical mean scores on the ninth grade administration of the SSHA.

The eleventh grade data show the groups in substantially the same relative position in terms of study habits mean scores as they were at the ninth grade testing. The final column in Table 6-4 gives the eleventh grade SSHA means after these means have been adjusted using covariance analysis to compensate for differences obtained in the ninth grade testing. These adjusted final means provide a better indication of the relative effectiveness of the two grouping treatments with respect to the study habits than do the unadjusted mean scores. It will be seen that the adjusted eleventh grade means for superior pupils in the two districts were very similar, being less than one raw score point apart. Slow pupils were also found to be very similar. Average pupils appear to have faired somewhat better in District R as indicated by the considerable differences in adjusted final means. In terms of the published norms, District A averagepupils were at the 35th percentile while District R pupils were at the 50th percentile (Brown and Holtzman, 1956).

Table 6-4

Mean SSHA Scores of Sample VIII Subgroups

| Group | 9th Grade | 11th Grade | 11th Grade Adj. for 9th Grade Dif. |
|---|---|---|---|
| District A, Superior | 31.69 | 34.37 | 31.89 |
| District R, Superior | 34.20 | 36.59 | 32.52 |
| District A, Average | 23.99 | 25.00 | 27.38 |
| District R, Average | 29.71[a] | 33.31[a] | 32.07[a] |
| District A, Slow | 24.38 | 20.86 | 22.99 |
| District R, Slow | 24.35 | 22.62 | 24.77 |

[a] This indicates a difference significant at the .01 level favoring the indicated group as compared with the comparable ability level in the other district.

### Discussion

It is interesting to note that in general the superior and average pupils in District R have somewhat higher initial study habit scores which will also be noted in comparisons of Sample IX pupils.

These differences reflect study habits developed prior to the start of the study and have no bearing upon the relative merits of the two grouping systems. The adjusted eleventh grade means, which reflect, at least in part, differences attributable to grouping show no differences except for average pupils, thus supporting the data obtained from the Sample VI analysis. We may conclude from the Sample VIII data that the high school pupil of average ability develops better study habits and attitudes in heterogeneous classes, probably because of the greater challenge or motivation provided him by the presence of superior pupils in his classes. On the other hand, there seems nothing to choose between the two grouping systems in terms of their effectiveness in developing better study habits in superior or slow pupils.

## Sample IX Data

### Introduction

Sample IX pupils were studied from the beginning of the ninth grade through the end of the twelfth grade. This sample was administered the Brown-Holtzman Survey of Study Habits and Attitudes on three occasions during the study: near the end of the ninth grade, near the end of the tenth grade and near the end of the twelfth grade.

The Sample IX data, along with the Sample VIII data are concerned primarily with study habits and attitudes developed during the high school years. These data are the best evidence available in this research concerning the relative merits of ability grouping versus random grouping for developing study habits during the aforementioned period.

### Results

The results of the three administrations of the SSHA for Sample IX were analyzed with two analyses of covariance. The first analysis adjusted the scores obtained at the end of the tenth grade for initial differences found in the ninth grade testing.

In view of the fact that District A employed ability grouping in mathematics and science only at the secondary level and the fact that prior to the start of this research project both districts had employed random grouping, one would expect differences on the SSHA administered in 1959 to reflect for the most part study habits differences that were developed prior to the outset of the study. It would seem likely that any major influence of ability grouping on study habits would be reflected in later tests; after pupils in District A had been exposed to the ability grouping system for a longer period of time. In comparing the mean scores for District A and District R pupils on the first administration of the SSHA (Table 6-5) we find fairly large differences between pupils of comparable ability. At all three ability levels these differences are in favor of the District R pupils with the differences being significant beyond the .01 level for superior pupils, significant at the .05 level for regular pupils and not statistically significant for slow pupils. Thus, we see that the District R pupils started the study with somewhat better study habits than District A pupils of comparable ability. For this reason, analysis of covariance provides the ideal tool for measuring performance of pupils in the two districts because it permits the later scores to be adjusted to compensate for these initial differences. It will be seen in Table 6-5 that the District R pupils tended to maintain their initial superiority throughout the three administrations of the Brown-Holtzman Survey of Study Habits and Attitudes. Generally, however, the differences became slightly smaller in the two later test administrations.

The first analysis of covariance carried out adjusts the scores obtained for Sample IX in February of their tenth grade year for differences that had been found on the initial test administered about one year earlier. All data needed to conduct this analysis were available for 219 Sample IX subjects. This analysis revealed that the adjusted mean scores for pupils of comparable ability in the ability grouped versus random grouped districts were not significantly different at any of the three ability levels.

As may be seen in Table 6-5, there were fairly large differences between pupils of different ability levels within each district. During the first year these differences were somewhat larger for District R pupils than for District A pupils, with the superior pupils obtaining the highest scores followed by the regular and slow pupils in that order. In general, within each district, the superior pupils obtained the highest mean scores in the three administrations of the Survey of Study Habits and Attitudes, the regular pupils were the next, and the slow pupils were lowest. The pattern of differences is substantially the same within each district.

The second analysis of covariance adjusted the study habits scores obtained in the twelfth grade for differences that were found in the tenth grade battery thus covering changes that occurred during the last two years of the study. All data needed to conduct the second covariance analysis were available for 214 pupils. The smallest subgroup was the District A superior with 28 subjects. Comparisons of the adjusted twelfth grade means of pupils of comparable ability in District A and District R showed no significant differences. It may be seen in Table 6-5 that the adjusted twelfth grade means for pupils of comparable ability in the two districts were very similar with slightly more than one raw score point separating the slow groups and even smaller amounts separating the means of the two average groups and the two superior groups.

In terms of gains in study habits scores over the period from the ninth grade test to the twelfth grade test, both superior groups made reasonably large gains. The average groups and slow groups, however, gained very little, earning virtually the same scores on the test in the twelfth grade that they had earned in the ninth grade.

## Discussion

The Sample IX results seem to suggest that neither the ability grouping system or the random grouping system as practiced by the districts involved in the Utah study make any unique contribution to the study habits of high school students. The almost identical adjusted twelfth grade means for the two superior groups obtained in the second covariance analysis indicate that, for this sample, ability grouping did little to pressure the bright students into developing better study habits as has often been suggested by advocates of ability grouping. The idea that average and slow pupils will develop better study habits in heterogeneous classes because of pressure placed upon them by superior classmates failed to gain support from the research data obtained from this sample.

Table 6-5

Mean SSHA Scores of Sample IX Subgroups

| Group | First Year (in Grade 9) | Second Year (in Grade 10) | Adjusted 2nd Year Means | Fourth Year (in Grade 12) | Adjusted[c] 4th Year Means |
|---|---|---|---|---|---|
| District A, Superior | 24.46 | 28.11 | 29.69 | 33.46 | 32.47 |
| District R, Superior | 32.57[a] | 34.47 | 30.28 | 37.90 | 32.76 |
| District A, Average | 24.51 | 24.68 | 26.23 | 25.73 | 27.78 |
| District R, Average | 28.94[b] | 28.75 | 27.14 | 28.97 | 27.43 |
| District A, Slow | 22.17 | 21.20 | 24.42 | 22.72 | 26.96 |
| District R, Slow | 24.76 | 24.53 | 25.90 | 24.37 | 25.89 |

[a] Significant at .01 level favoring group indicated.
[b] Significant at .05 level favoring group indicated.
[c] Adjusted for 10th grade differences using covariance analysis.

## Ability Level Differences in Study Habit Scores

### Introduction

Although not directly pertinent to the main goal of the research project, comparisons between pupils at different ability levels in each district were also made in the analyses of variance and covariance that were carried out primarily to compare pupils of comparable ability in Districts A and R. These comparisons are summarized in Table 6-6. All statistically significant differences favored the group of higher ability. In all, 39 comparisons were made in each district between pupils at different ability levels: 13 comparing superior versus average pupils, 13 comparing superior versus slow pupils, and 13 comparing average versus slow pupils.

### Results

In District A, of the 13 comparisons of mean study habit scores and subscores between superior and average pupils, 11 were statistically significant, 10 of these beyond the .01 level. Comparisons of superior versus slow pupils in District A yielded 12 statistically significant differences out of 13, comparisons with 11 beyond the .01 level. Comparisons between average and slow pupils in District A yielded only three statistically significant differences out of 13 with two of these significant beyond the .01 level.

These findings are generally supported by the District R data. Comparisons of mean scores of superior versus average pupils in District R yielded 10 statistically significant differences out of 13 comparisons with seven of these beyond the .01 level. All 13 comparisons between the mean scores of superior and slow pupils in District R were statistically significant with 12 beyond the .01 level. Comparisons of average versus slow pupils in District R yielded eight statistically significant comparisons with five of these beyond the .01 level.

In reviewing these data, we find that 46 of the 52 comparisons between superior pupils and either average or slow pupils, showed that superior pupils had significantly higher mean study methods scores. The differences between average and slow pupils were somewhat less marked, with 11 of the 26 comparisons made in the two districts being statistically significant. This, however, is still a sizeable number of significant differences and seems to indicate a fairly consistent relationship between ability level, as measured by a composite achievement score, and study methods. This relationship may well have been less if a more direct measure of ability such as an intelligence test had been employed by District A in their grouping procedures instead of an achievement battery.

These data, by showing the consistently higher study habits scores of superior pupils as compared with average or slow pupils, further refutes the opinion of many educators that bright pupils do not learn to study because of the lack of challenge for them in today's schools.

## Table 6-6

## Differences in Study Habits Scores Between Pupils of Different Ability Levels in Each District[a]

| | Sample IV 7th Grade CSMS Study | | | | Sample VI 9th Grade CSMS Study | | | | Sample VIII 9th Gr. 11th Gr. | | Sample IX 9th 10th 12th | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Levels | Att. | Mech. | Plan. | Total | Att. | Mech. | Plan. | Total | SSHA | SSHA | SSHA | SSHA | SSHA |
| Superior vs. Aver. District A | .01[b] | .01 | .05 | .01 | .01 | .01 | .01 | .01 | .01 | .01 | NS | NS | .01 |
| Superior vs. Slow District A | .01 | .01 | .05 | .01 | .01 | .01 | .01 | .01 | .01 | .01 | NS | .01 | .01 |
| Average vs. Slow District A | NS | .01 | NS | .05 | NS | .01 | NS | NS | NS | NS | NS | NS | NS |
| Superior vs. Aver. District R | .01 | .01 | .01 | .01 | .05 | .01 | NS | .05 | .05 | NS | NS | .01 | .01 |
| Superior vs. Slow District R | .01 | .01 | .01 | .01 | .01 | .01 | .05 | .01 | .01 | .01 | .01 | .01 | .01 |
| Average vs. Slow District R | .01 | .05 | NS | .05 | .01 | .01 | NS | .01 | .05 | .01 | NS | NS | NS |

[a] In all cases where significant differences between means were found, the difference was in favor of the pupils of higher ability.

[b] ".01" indicates a difference significant at the .01 level; ".05"-- a difference at the .05 level and "NS"--no significant difference.

## SUMMARY AND CONCLUSIONS

### Elementary

The Sample IV results on the CSMS showed all of the total score differences and seven of the nine subtest differences between comparable groups in Districts A and R to be statistically significant. Differences were generally largest between pupils of superior ability. These results lead to the conclusion that pupils in a random grouping situation consistently develop better study habits during the elementary school years than pupils in an ability grouping situation.

### Secondary

Sample VI average pupils in District R received consistently higher scores on the CSMS than comparable District A pupils. No significant between-treatment differences were present for slow or superior pupils. Sample VIII average pupils in District R received higher SSHA scores. No significant between-treatment differences emerged from the Sample IX data. If any advantage accrues from heterogeneous grouping at the secondary level, it is for average pupils, who made significantly greater gains in study habits in Samples VI and VIII. Although this difference was not found in Sample IX, the larger numbers of subjects in the other samples would still justify this conclusion.

### Level Differences

In comparing the study habits scores of pupils of different ability levels within each treatment a marked relationship between study habits and ability level was apparent. Of the 52 differences comparing superior pupils with average and slow pupils, 46 were statistically significant as were 11 of the 26 comparisons made between average and slow pupils. All significant differences favored the group of higher ability. These results support the conclusion that superior pupils generally develop superior study habits in both grouping treatments. This evidence counters the frequently heard criticism that superior pupils do not learn to study in the public schools.

## REFERENCES

Brown, W. F., & Holtzman, W. H. SSHA Manual, Revised. New York: Psychological Corporation, 1956.

Carter, H. D. Manual, California Study Methods Survey. Los Angeles: California Test Bureau, 1958.

## Chapter 7

# SOCIOMETRIC CHOICE

by Walter R. Borg and David Brinks[1]

## INTRODUCTION

People are an important part of the environment to which individuals react in the processes of developing and learning. These people may be in groups; families, classrooms, communities, social classes or nations. It is the individual, however, along with a number of other individuals, that comprise the group. In other words, the individual and the group are interdependent. Wrightstone, Justman, and Robbins (1956, p. 199) have asserted that "The development of one's personality is a product of the interaction of the individual with other people." According to Otto (1954) it is in these person-to-person relationships in various sizes and types of groups that many social needs are met and personality development has its roots.

The need to belong to a group is a powerful force. Some psychologists claim that this need is an innate human need while others maintain that it is culturally developed. All agree that, whatever its origin, the need to belong furnishes strong motivation for behavior, for in our society almost every individual attempts to become a member of some group. Cunningham, et. al. (1951, p. 108) state that "It is important for the development of most boys and girls of school age that they find belonging in groups of their peers." Dineen and Gary (1955, p. 358) expressed their agreement with this concept in maintaining that,

> Chief among the factors necessary for developing a full life in a heterogeneous group such as a class of children is a satisfactory friendship pattern. When a child is with others who accept him and respond to him, that is others with whom he wants to associate, he can contribute more and function better in the group.

Social adjustment has become of major importance in education. The alert teacher is aware of the need of children to establish and maintain emotional and social stability. The relationship which exists between social adjustment and academic achievement has been shown in several recent studies (Buswell, 1953; Thorpe, 1955; Carew, 1957; and Ryan and Davie, 1958).

In ability grouping, the matter of social relationships and adjustment has been of primary concern to critics. One of the criticisms most often directed at ability grouping in the public schools is that the children will suffer social maladjustment as a result of being placed in a classroom limited to fellow students of similar intellectual caliber. The attitude of many people is revealed in Brueckner and Bond's statement (1955, p. 50) that "Current educational practices that segregate the good, average, and poor achievers within the classroom tend to exaggerate the child's social adjustment." Luchins and Luchins (1948) reported finding that grouping caused slow students to develop feelings of inferiority and resulted in bright students becoming snobs. Such a system results in an artificial social situation which stigmatizes slow students, according to Hamalainen (1950).

---

[1] The parts of this chapter dealing with sociometric perception and sociometric stability are based upon Dr. Brinks' doctoral research carried out under Dr. Borg's supervision.

Many of those who advocate ability grouping claim that the slow child has an opportunity to be recognized socially in such groups. He is denied such recognition in randomly grouped classes where he must compete with more able students. Studies by Justman (1953), Mann (1957), and Goldberg (1958), however, showed little difference in the social and personal attitudes of students resulting from the two plans for grouping students.

Although many studies have explored achievement in homogeneous grouping, only a very limited amount of research has investigated directly the social structure and relationships in classrooms using ability grouped and randomly grouped systems. Because social facility contributes so extensively to academic performance it would appear that a thorough investigation of peer relationships in different grouping situations would be of considerable value.

The primary purpose of this phase of the Utah study was to investigate the effects of homogeneous versus heterogeneous grouping practices on the sociometric structure of elementary school classrooms. This represented an attempt to determine if systematic differences exist which may be attributable to the type of grouping being practiced. Sociometric status patterns, accuracy of perceived status, and the long-range stability of both was studied in three replications, using the same subjects, Sample IV, at the fourth, fifth, and sixth-grade levels in three successive academic years. In addition, sociometric status patterns only were studied using Sample VI pupils during their sixth grade year.

Sociometric status is defined as the total peer choice score or the total number of times a pupil was selected on a sociometric choice measure by his classmates as a preferred companion on any one of the three specified activities or situations used as criteria. By accuracy of perceived status is meant the accuracy of the individual student's estimation or "guess" of who would choose him as a preferred companion for the activities or situations which were specified in the sociometric measure. Accuracy of perceived status was determined by comparing a pupil's guesses of who would choose him in each of the criteria with those who actually did choose him.

Mobility of sociometric status may be defined as movement from one sociometric classification[1] to another sociometric classification on measures obtained in consecutive years. Mobility of status and consistency of perceived status from year to year over the three year period were studied in this project. This was done in an attempt to determine whether the social relationships measured were transitory or relatively stable. It was felt that replicating the study could also give some indication of the degree to which increasing exposure to the two grouping treatments and increasing age and maturity, with its accompanying changes in personality characteristics and interests, might affect one's social relationships. Are status and acceptance based on stable personality traits or, as Staker (1948) has suggested, are there different traits which are important at different age levels? It was felt important to verify by replication the findings of the first year, as this yielded valuable information pertaining to the reliability of the sociometric measurement, as well as providing a cross check on the initial findings.

---

[1] The three sociometric classifications used in this project were star, regular, and neglectee-isolate.

Gronlund and Whitney (1955) have suggested that social acceptance pervades all aspects of one's life sphere. This would mean that those factors that contribute to one's acceptance by others operate at home, on the job, at school, in informal contacts, or in any situation in which a person may become involved. This idea was supported by Symonds and Jackson (1935). Other researchers (Bronfenbrenner, 1945; Biehler, 1954; Gronlund, 1959; and Bjerstedt, 1956) presented evidence showing the generality of social status over criteria. Their findings have suggested that one who obtained a high sociometric score on one criterion would probably receive relatively high scores on other criteria. An attempt was made in the Utah study to determine if accuracy of perceived status was consistently stable over criteria, that is, from one criterion to another.

Another objective of this research was to determine factors related to sociometric status and accuracy of perceived status. The relationship between sociometric status and accuracy of perceived status was studied, as well as the relationship between sociometric status and other variables. These variables included mental ability, academic achievement, attitudes toward school, teacher, and peers, and self concept. An attempt was made to determine the effect of the different grouping situations on each of these relationships.

The design of the research provided for the investigation of the effects of three different conditions on measures of sociometric status and accuracy of perceived status. These conditions were (1) the grouping treatment (ability grouped or random grouped), (2) ability level, (based on academic achievement) and (3) sex. Comparisons were made between the two districts, between the various ability levels, and between sexes on the aforementioned dependent variables.

## WORKING HYPOTHESES

On the basis of a comprehensive review of related research, the following working hypotheses were established regarding the effects of grouping treatment, ability level, and sex on the sociometric variables.

### Sociometric status patterns

A number of previous studies have shown relationships between sociometric choice scores and achievement or ability in heterogeneous classrooms. When one considers the ability grouping situation, however, where achievement variability is somewhat reduced, it appears probable that the effect upon superior pupils would be to reduce their chances of gaining social recognition. Two changes in the classroom social situation seem possible for average and slow pupils when placed in an ability grouped situation. One of the possible effects of ability grouping could be the development of a leadership vacuum in the average and slow groups resulting in few if any recognized social leaders or stars[1] in these groups. The second alternative would be that, with the superior pupils removed, pupils in average and slow groups who had not previously been recognized in the peer structure would emerge as social leaders or stars. If the first of the two aforementioned alternatives occurred, the over-all proportion of stars would be considerably less in District A than in District R. On the other hand, if the latter situation occurred, the number of stars in District A would be equal or perhaps greater than in District R, as fewer superior pupils would probably attain the status of stars in the ability grouped class but greater numbers of both average and slow pupils would reach this level of acceptance.

---

[1] See page 164 for definition of the sociometric catagories.

Working hypotheses concerned with sociometric choice patterns in the two districts involved in the Utah study were based upon the latter expectation. Thus, it was hypothesized that between-treatment comparisons would reveal fewer stars among District A superior pupils but greater numbers of stars among both average and slow pupils in District A. Furthermore, it was hypothesized that the combined distribution of pupils at all ability levels into the three sociometric categories of "stars," "regulars" and "neglectees-isolates," would not be significantly different for the two districts.

Regarding level differences within each treatment, it was hypothesized that in the random grouped situation, where previous studies have demonstrated a relationship between status and ability, superior pupils would obtain the more favorably sociometric status patterns (more stars and fewer neglectee-isolates) followed by average and slow pupils in that order. In District A, where ability differences within the classroom were somewhat less, it was hypothesized that sociometric choice patterns for pupils at the three ability levels would not be significantly different. That is, superior, average, and slow pupils would all have about the same distribution of stars, regulars, and neglectee-isolates.

It was hypothesized that no significant sex differences would occur in the sociometric choice patterns of boys and girls of comparable ability.

### Accuracy of Perceived Status

It appeared likely that the pupil's ability to predict accurately who among his peers would choose him in the various sociometric categories would by largely a function of intelligence. Thus, it was hypothesized that no significant between-treatment differences would be found in comparing the accuracy of perceived status for District A and District R pupils of comparable ability.

With regard to level differences, it was hypothesized that in both districts comparisons between pupils of different ability levels would reveal pupils of higher ability to be more accurate in their perceptions than pupils of lower ability.

It was hypothesized that in comparisons between boys and girls of comparable ability, no significant sex differences would emerge in accuracy of perceived status.

### Mobility of Sociometric Status

The degree to which pupils in the two treatments changed from one to another sociometric classification from year to year appeared likely to provide some valuable insights into the relative effect of the two grouping treatments upon the classroom social structure. Having already hypothesized that average and slow pupils would tend to attain more favorable sociometric status in the ability grouped situation, it appeared logical to hypothesize that the overall patterns for average and slow pupils would tend to improve from year to year in the District A classrooms while staying about the same in District R classrooms. This hypothesis was based on the assumption that many average and slow pupils might require a certain amount of learning and experience before they would be able to move out of the background and assume positions of social leadership in ability grouped classes. Thus, although it might be expected that some average and slow pupils would move up into positions of social leadership very quickly, it seemed likely that many who had potential for social leadership would require a longer experience before emerging as stars. This lag could result in a temporary leadership vacuum in average and slow sections in District A.

In terms of level or within-treatment differences in social mobility patterns, it was hypothesized that no significant pattern of change would emerge in District R. It was hypothesized, however, that in District A an upward mobility pattern would emerge for slow and average pupils, but that no marked status changes would occur for superior pupils over the three years of the study.

Regarding sex differences, it was hypothesized that no significant differences in social mobility patterns would occur between boys and girls of the same ability level and within the same grouping treatment.

### Accuracy of Perceived Status over a Three Year Period

Regarding the long range consistency of pupils in their ability to accurately perceive their status over the three year period of the study, it was hypothesized that consistency would be correlated with ability level, but that no significant differences would occur between pupils of comparable ability in the two treatments, or between the sexes.

### Relationships Between Sociometric Variables and Other Measures

A final goal of this phase of the project was to determine whether significant relationships existed between sociometric status, accuracy of perceived status, and a number of other variables for which data were collected in other phases of the Utah study. It was hypothesized that sociometric status and accuracy of perceived status would both correlate positively with ability and achievement for District R pupils but that these variables would not be significantly correlated for District A pupils. It was hypothesized that a positive relationship would exist in both districts between sociometric status and favorable attitudes towards the school environment. Positive relationships were also hypothesized between sociometric status, accuracy of perceived status, and self concept variables as measured by the Index of Adjustment and Values.

## PREVIOUS RESEARCH RELATED TO GROUPING AND SOCIOMETRIC CHOICE

### Sociometry

"Interpersonal relations are the life-blood of society, and creators of mind and self . . . therefore, an understanding of these relations is essential to societal adjustments." (McKinney, 1947, p. 348) One device for studying human relations is known as sociometry. The term "sociometry," derived from Latin, means social or companion measurement. Sociometry explores the structure of the group and the place of the individual within that group. It measures the social acceptability of a member of a group relative to the social acceptability of other members of the same group. Sociometry measures not only the direction of interpersonal relations, but also the intensity of them. It illustrates relationships as they actually are, not as they appear to be.

Some teachers will argue that they can identify the social position of pupils in their classrooms without using sociometric measures. Usually these teachers will be surprised when they use sociometric methods. As research has shown, and as Ohlsen (1955) pointed out, the teacher frequently sees group relationships differently than do the pupils. These incorrect perceptions with which the teacher may attempt to work makes her efforts to improve the pupils' ability to live and work in groups relatively ineffective. Sociometric methods, providing approaches for obtaining peers' reactions to individual pupils, form a sound basis for therapeutic action and readjustment.

There are myriad factors which influence pupils in their choices of companions. Some of these factors include: the type of group, chronological age, mental age and intelligent quotient, home background, and personality traits (Blanchard, 1947).

## Sociometric Terms

The term "sociometric status" refers to the number of choices that an individual receives on a sociometric test. Gronlund (1959) preferred the use of this term or "group status" rather than "social status," since the latter is frequently used in other situations, such as status in a community.

A "star" is defined as an individual who receives a large number of choices on a sociometric test, more than could be expected by chance alone. One who receives relatively few choices on such a test, fewer than could be expected by chance, is known as a "neglectee," while those falling between stars and the neglectees are "regulars." An "isolate" is one who receives no choices, while a "rejectee" is one who receives negative choices, if negative choices have been obtained.

A situation in which two individuals choose each other on the same criterion is known as "mutual choice." It is a reciprocated choice. A number of individuals who choose each other on the same criterion, but give relatively few choices to individuals outside their closely knit group, is known as a "Sociometric clique." "Wociometric cleavage" results when there is a lack of sociometric choices between two or more sub-groups.

## Ability and Sociometric Status

Jennings (1950) studied the relationships between mental ability and sociometric acceptance. In a longitudinal study involving over 400 girls, she found virtually no relationship between the extent to which a group member was chosen by others and intelligence. Challman (1932) discovered that similarities in mental ages had no effect on the friendship choices of either boys or girls. These findings were substantiated by Cunningham, et al. (1951), who found little significant correlation between acceptance and intelligence quotient.

Routh (1960) utilized data which had been obtained from sixth grade public school children on three sociometric questions. He found that in terms of average intelligence, there was no statistically significant difference between the over-chosen (over one standard deviation above the mean) and the under-chosen (lower than one standard deviation below the mean) groups. Both of these groups scored above the average-chosen group, suggesting that greater intelligence could contribute either to higher or lower sociometric status. Although Routh claims statistical support for his findings, none were presented.

Ryan and Davie (1958), using a sample of 326 subjects, the entire high school population of a suburban town, found a positive, but not statistically significant relationship between classroom grades and social acceptance. They report that verbal aptitude, as measured by the American Council on Education Psychological Examination, did not correlate significantly with social acceptance.

Gallagher (1958), in a well designed study, found that pupils with higher levels of intellectual ability tended to receive more choices than those with lower levels of ability. He used 355 children from three classrooms at each grade level, grades two through five, from eight different schools. Assessment devices included the California Test of Mental Maturity, the Otis Test of Mental Ability, and a five choice sociometric test. On the basis of his findings he was forced to reject his hypothesis that children of similar intellectual levels have a tendency to choose each other as friends more frequently than they will choose children of different intellectual levels. Barbe's work (1954) also demonstrated that children in heterogeneously grouped classrooms would select as their friends children from different levels of intelligence. Slow children tended to make the majority of their friendship choices from the high average group, while bright children chose their friends primarily from the bright group. Bright children were chosen by the average group as friends far more frequently than were slow children. Each group, however, listed some friends from all levels. This suggests that intelligence may not be as important as has been assumed in some cases, although Furfey (1927) and Jenkins (1931) both reported a slight tendency for children to choose friends of approximately the same intelligence.

Taylor (1952), who also used the California Test of Mental Maturity, found a significant positive relationship between mental age and social acceptance. Thorpe (1955) found a significant positive correlation between intelligence and sociometric status in a study of 900 children from 34 English secondary schools.

Bonney (1942) found some relationship between intelligence and mutual friendship scores. He concluded that even though higher degrees of brightness were definitely associated with the ability to win friends, intellectual brightness was far from being any kind of a guarantee of social competence. He supported these conclusions by the findings of later studies (1943; 1944). Although he again found positive relationships between intelligence and academic achievement and measures of social success, he cautioned against assuming the attainment of social skills to be the natural sequence of intellectual brightness or the mastery of subject matter. In still another study Bonney (1946) found general intelligence to be somewhat more related to the formation of mutual friendships than was academic achievement. The evidence indicated that the degree of academic proficiency was not a discriminating factor in determining friendships.

Morgan (1946), in studying all 274 children enrolled in grades five, six, seven, and eight of the Seneca, Illinois, public schools, found that the academically successful children were those most favored socially. Those who scored low in achievement were more often rejected and were unfavorably rated on social reputation.

Buswell (1953), in an attempt to determine if children who were accepted by their peers differed in achievement from those who were rejected, found a very definite difference between the two extremes. This difference was apparent in the upper elementary grades, but not at the kindergarten level. She found, however, that if the factor of intelligence were controlled, the differences no longer existed. Apparently the influence of intellectual ability on social acceptance increased with age, at least through the elementary grades.

Heber (1956) and Davis (1957) both found "popular" children to be more intelligent than those low in sociometric status. Heber presented evidence which suggested that a low Intelligence Quotient has a greater decremental effect on sociometric status than the incremental effect of a high Intelligence Quotient.

Laughlin (1954) conducted a two-year study of peer status. The first year 630 sixth-grade children in the Port Arthur, Texas, public schools were studied. The following year the same sociometric instruments were used on the same children in the seventh grade. There were 525 subjects common to both grades.

The correlation coefficients between group social acceptance scores, mental age and I.Q.'s for each grade level are shown in Table 7-1. The relationship between group social acceptance and mental ability was found to be consistent in both years suggesting a true but slight relationship. Mental ability had a higher correlation with peer status than did academic achievement. Children with high I.Q.'s and high academic achievement scores were, with few exceptions, well accepted by their group. However, average children also were frequently popular. Pupils who were especially disliked were seldom those with the lowest I.Q.'s.

A study of the adjustment of 30 gifted children in the regular classroom by Gallagher and Crowder (1957) found that, despite many deviations from the group norm, the gifted children as a whole showed superior social popularity. On sociometric measurement, 80 per cent were above the average, with 53 per cent in the top quartile. This finding was confirmed in another study by Gallagher (1958) in which he studied 29 boys and 25 girls in heterogeneous classes, grades two through five, who had Stanford-Binet Intelligence Quotients of 150 or over. In each class where a highly gifted child was found, sociometric choices were obtained from all pupils. He found the gifted children to be more socially accepted by their peers than were the children of average intelligence in the classroom, although the popularity of the gifted group decreased as their intelligence reached an extremely high level. He found no support for the hypothesis that gifted children would choose their friends near their own intellectual level.

Evidence reported by other researchers (Grace and Booth, 1958; Bedoian, 1954; and Nelson and Carlson, 1945) supports the findings already presented.

It has been shown that gifted children as a group are above average in social development and have fewer problems of personality adjustment than do average or lower ability children. A low but consistently positive relationship has been found between intelligence and status as measured by sociometric techniques. There are exceptions, of course. Many gifted children experience serious difficulties in meeting life situations. Social and emotional maturity are not inherent in giftedness. The high intelligence of these persons gives them insights that may be helpful in solving problems, but may also be a source of deep sensitivity that compels them to face problems not ordinarily met by other children.

Table 7-1

Correlation Coefficients Between Group Social Acceptance and Mental Ability[a]

| | Sixth Grade | | Seventh Grade | |
|---|---|---|---|---|
| | Mental Age | I.Q. | Mental Age | I.Q. |
| Group Social Acceptance | .327 | .307 | .280 | .265 |

[a] Taken from Laughlin, (1954).

## Social Aspects of Ability Grouping

Ability grouping has probably been criticized more frequently in regard to its social aspects than it has in regard to academic achievement. Charges that ability grouping causes social maladjustment are typical of the tirades loosed in opinion articles by Tonsor (1953), Bettelheim (1958), Cavanaugh (1960), and Ginzberg (1961). Lawson (1957), conversely, contended that homogeneous grouping would not cause maladjustment in the child. He felt that there is a natural tendency for children to gravitate toward contacts with other children of similar intellectual levels. When a child is with others with whom he cannot compete he meets constant failure that is apt to cause maladjustment. Goldberg (1958) also felt that placing youngsters into narrow-range ability groups would not bring about poor social or personal adjustment. McGaughy (1930), in listing opinions for and against ability grouping, pointed out that students with similar rates of speed are happier together and enjoy school life more, but also that children in slow groups are marked as "dullards" or "dumbbells."

It is readily apparent that there is no uniformity of opinion regarding the effect of ability grouping on the social development of the participants. The sources cited above, being primarily opinion, may or may not find support in research findings.

There is evidence available that indicates the claim that students are more likely to develop undesirable social attitudes in ability grouped sections than in ungrouped sections is not well founded. Studies by Goldberg, Passow and Lorge (1958), Goldworth (1959), Rochfort (1959), and Bell (1959) showed little difference in the social or personal attitudes of students growing out of the two plans for grouping. Barbe (1955), on the basis of a 76.9 per cent return of a questionnaire mailed to 456 students of special high school classes for gifted children, found that the majority of the respondents reported that having been members of these special classes aided them in making a good adjustment. This contradicted the usual complaint about homogeneous grouping to the effect that such children have a more difficult time adjusting to other groups. Such data, however, must be interpreted with caution because of the possibility that the responding group represents a biased sample.

Jackson (1943), working with third and fifth grade pupils, reported findings generally favorable to ability grouping. Not only were the schools bothered with only minor behavior problems, (no mention was made of the extent of behavior problems earlier) but these schools found that in ability grouped sections truancy almost disappeared.

Martinson's (1959) study of children, that we have discussed in previous chapters presents some interesting evidence on sociometric choices. Her experimental group consisted of 929 boys and girls from every type of public school designed to represent the entire school population of California. The mean Intelligence Quotient of the sample, based on the revised Stanford-Binet scale, was 142.6. The subjects were matched by pairs with a control group on the basis on chronological age, Intelligence Quotient, sex, and socioeconomic status. The programs in which the subjects participated were planned within the general areas of enrichment within the classroom, acceleration, and special groupings. In these areas a total of 17 programs at various grade levels were evaluated.

It was found that first-graders in the experimental programs showed highly significant gains in friendship choices, while the control gifted subjects showed no change, either positive or negative. Again at the fifth and sixth-grade levels, the total experimental group of 191 subjects showed highly significant gains in social status on the basis of responses by their peers in regular classroom situations, while their matched controls showed no change. The author concluded that "Fears that special planning necessarily penalizes the gifted child socially are unfounded." (1959, p. 342) Apparently the gifted child's classmates did not regard him as a less desirable friend because special provisions were made for him. Goldberg (1960, p. 51) supported these findings in her statement that "There is no evidence to support the notion that grouping has adverse effects on the social or personal attitudes or behavior of children."

Drew's study (1962), was concerned with the effects of homogeneous and heterogeneous grouping on the performance of superior, average and slow pupils in ninth grade English classes. Her data provide some valuable evidence on the effects of grouping upon sociometric status. In her study, six three-choice sociometric questions were used. Two of these questions required pupil choices on social variables while the other four were concerned with intellectual variables. Pupil choices were analyzed question by question and also composite peer choice scores in the social and intellectual categories were calculated. The pupil's choices were limited to individuals in his English class, which places considerable restriction on choices in the departmentalized junior high school situation.

In the heterogeneous classes, Drews found that for each of the six sociometric criteria superior students were chosen more frequently than average students and slow students were least often selected. Although marked differences occurred in both the social and intellectual categories, the differences, as might be expected, were greater in the intellectual area. Drews also compared the distributions of sociometric choice scores obtained by superior, average, and slow pupils in her homogeneous and heterogeneous classes. Comparisons of mean scores in the two types of classes would, of course, be spurious because in the homogeneous class the pupil has no choice but to nominate persons of the same ability level. The distribution of choice scores, however, can be meaningful because if a leadership vacuum exists in the average and slow homogeneous classes as has sometimes been suggested, none of the average and slow pupils in homogeneous classes would receive enough sociometric choice nominations to be classified as "stars." The percentage of students receiving different numbers of nominations in Drews' study are summarized in Table 7-2. It will be seen that although a larger percentage of superior pupils received high numbers of nominations in both grouping systems, the differences were much greater in the heterogeneous classes.

As might be expected, the differences on the intellectual choice questions were much more pronounced than on the social choice questions. For example, if we take a sociometric score of 12 as the cut-off point, we find that 54 per cent of the superior pupils in heterogeneous groups obtained more than 12 nominations while no slow pupils in the heterogeneous groups reached this level. In contrast, 30 per cent of the superior pupils in homogeneous groups received more than 12 nominations while 28 per cent of homogeneously grouped slow pupils reached this point.

On the social questions, approximately 33 per cent of the superior pupils in heterogeneous groups received more than 12 nominations with no slow pupils receiving this number. In homogeneous groups, 30 per cent of the superior pupils, and 18 per cent of the slow pupils, received 12 or more nominations. These data would indicate that although a smaller proportion of slow than superior pupils in either treatment gained positions of social recognition, the slow pupil had almost no chance

## Table 7-2[a]

Per Cent of Students in Achievement Groups Receiving Intellectual and Social Nominations in Homogeneous and Heterogeneous Classes

| Class and Group | N | Number of Nominations Received<br>30+ | 21 - 29 | 13 - 20 | 9 - 12 | 5 - 8 | 1 - 4 | 0 |
|---|---|---|---|---|---|---|---|---|
| | | | | | Intellectual Questions | | | |
| Homogeneous | | | | | | | | |
| Superior | 101 | 10.89% | 9.90% | 9.90% | 8.91% | 18.81% | 39.60% | 1.98% |
| Average | 146 | 9.59 | 7.53 | 10.27 | 13.70 | 16.44 | 37.67 | 4.79 |
| Slow | 78 | 5.13 | 5.13 | 17.95 | 20.53 | 23.08 | 25.64 | 2.56 |
| Heterogeneous | | | | | | | | |
| Superior | 42 | 35.71 | 14.29 | 4.76 | 7.14 | 21.43 | 14.29 | 2.38 |
| Average | 171 | 6.43 | 4.09 | 12.28 | 9.36 | 17.54 | 40.94 | 9.36 |
| Slow | 27 | 0 | 0 | 0 | 0 | 3.70 | 70.37 | 25.93 |
| | | | | | Social Questions | | | |
| Homogeneous | | | | | | | | |
| Superior | 101 | 0% | 8.91% | 20.79% | 14.85% | 16.83% | 28.71% | 9.90% |
| Average | 146 | 2.05 | 4.11 | 17.12 | 14.38 | 26.03 | 30.82 | 5.48 |
| Slow | 78 | 0 | 2.56 | 15.38 | 10.26 | 33.33 | 33.33 | 5.13 |
| Heterogeneous | | | | | | | | |
| Superior | 42 | 2.38 | 16.67 | 14.29 | 11.90 | 14.29 | 30.95 | 9.52 |
| Average | 171 | 1.17 | 4.09 | 18.13 | 10.53 | 23.39 | 35.09 | 7.60 |
| Slow | 27 | 0 | 0 | 0 | 18.52 | 14.81 | 51.85 | 14.81 |

[a] Taken from Drews (1962), page P-13.

to gain such status in the heterogeneous classes, while his chances were nearly as good as the superior pupil in the homogeneous classes. If Drews' study had extended for more than one year, it seems likely that the distributions of sociometric choice scores in the homogeneous classes would have become even more similar as slow pupils gained more experience in positions of social leadership. Many psychologists have stressed the importance of peer acceptance during the adolescence period. In view of this, Drews' evidence would seem to provide strong support for ability grouping at the junior high school level.

Mann's research (1957) concluded that there was little real friendship between superior and average or below average students in heterogeneously grouped classrooms. In this study 67 gifted pupils in a total sample of 280 fifth and sixth graders were placed in regular classes for half of their work and in special workshop rooms with other gifted children for the other half. On the basis of sociometric choices, there were more acceptances and rejections of the gifted by the gifted than by the typical children. Typical children, as a group, tended to accept and reject more typical children than workshop children. These findings suggest that friendships were based more on like mental age than on like chronological age.

It appears to be a safe assumption that students will select as friends those persons with whom they associate more, especially at the elementary school level. With this in mind one might speculate as to whether the sociometric results of Mann's study were due to similar mental ability or to the greater contact among the gifted children than between the gifted and typical groups. This increased contact would serve to increase the number of sociometric choices regardless of the basis used for separating the children into groups, thus casting serious doubt on the validity of Mann's conclusions.

Luchins and Luchins (1948) individually interviewed 190 fourth, fifth and sixth grade children in an attempt to determine the attitudes of children in ability grouped schools. Their findings supported some of the critics of ability grouping. Apparently all of the children were aware of the grouping system and the resulting caste system it helped create in the school. The lower groups appeared to feel inferior and ostracized and were aware of the stigma attached to membership in their class. Since the subjects in Luchins and Luchins' study were not compared with any controls, it is impossible to conclude definitely that the findings were attributable to the grouping situation. While this appeared to be the case, the study was not so designed to allow this conclusion. Students' feelings, like those reported in Luchins and Luchins' study, appear to be evident to school administrators. A survey by Saville (1961) revealed that a number of school officials considered the "stigma" attached to pupils involved in ability grouping to be a major problem.

The literature concerning the social aspects of ability grouping includes at least some evidence to support any stand one might care to take. Most of the research studies concluded that the grouping programs had no adverse effect on the social or personal adjustment of the students. Some evidence was presented which indicated an increase in improved social relationship, while other evidence supported many critics of ability grouping by indicating that the development of undesirable attitudes resulted from such grouping programs. The inconclusive nature of these studies points to the great need for additional research in this area.

## PROCEDURE

### Sample

Table 7-3 gives the number of Sample IV and Sample VI pupils by district, ability level, and sex who participated in this phase of the Utah study. Sample VI pupils were administered the sociometric measures only once--in the sixth grade--and therefore, this sample was used only in the study of sociometric status patterns. Sample IV pupils were measured in the fourth, fifth, and sixth grades, and data from this sample were used to study sociometric status patterns, accuracy of sociometric perception, and stability of sociometric status.

### Measures Used

A near-sociometric measure of the usual partial-rank-order type was developed for use in this study. Three criteria were employed, each with five positive choices. Each subject was asked to indicate by placing a check mark beside names on a roster of his classmates those whom he preferred under each criterion. The three criteria were: (1) the five children in the class who are your best friends; (2) the five children with whom you prefer to study; and (3) the five children who you would most like to have with you if you were transferred to another classroom.

A perceptual element was added to this device. This consisted of having each student estimate or "guess" who would choose him as a preferred companion on each of the three criteria. This was done by having him place check marks behind the names on the same rosters on which he had indicated his preferences.

Measures from other parts of the study with which the sociometric data were compared included the *Bills Index of Adjustment and Values*, which was used to measure the self-concepts held by the students, and the *Sequential Tests of Educational Progress*, that were used to obtain a measure of achievement. The *California Test of Mental Maturity* provided an indication of mental ability other than that available in terms of achievement test scores. The *USU School Inventory*, a locally developed attitude scale, was used in an attempt to assess attitudes of students toward school, teachers, and peers.

Each pupil's sociometric status score was obtained by summing the choices he received from others on the three criteria. After obtaining a sociometric choice score for each pupil, the pupils were classified as stars, regulars, neglectees, or isolates based upon choice limits established by Bronfenbrenner (1945). Category limits using this system were as follows:

"stars" __________ received twenty-two or more nominations
"regulars" ________ received ten to twenty-one nominations
"neglectees" ______ received one to nine nominations
"isolates" ________ received no nominations

For purposes of the analysis, however, neglectees and isolates were combined. An extremely small number of isolates were obtained in the three-criterion five-choice sociometric measure.

Data needed to estimate the pupil's accuracy of perceived status were obtained by having each student guess who would choose him as a preferred companion in each of the three choice situations. Comparing the student's guesses of who would choose him with those who actually did choose him gave an accuracy score. The

Table 7-3

Numbers of Subjects in Each Subgroup for Each Replication of Sociometric Study[a]

| Group | 1958-59 4th Grade (Sample IV) | 1959-60 5th Grade (Sample IV) | 1960-61 6th Grade (Sample IV) | 1958-59 6th Grade (Sample VI) |
|---|---|---|---|---|
| District R: | | | | |
| 1. Superior boys | 121 | 121 | 165 | 124 |
| 2. Average boys | 79 | 110 | 116 | 97 |
| 3. Slow boys | 59 | 80 | 91 | 65 |
| 4. Superior girls | 136 | 136 | 163 | 141 |
| 5. Average girls | 72 | 113 | 122 | 75 |
| 6. Slow girls | 29 | 35 | 45 | 35 |
| Totals | 496 | 595 | 702 | 537 |
| District A: | | | | |
| 1. Superior boys | 45 | 60 | 50 | 69 |
| 2. Average boys | 56 | 129 | 88 | 73 |
| 3. Slow boys | 53 | 59 | 37 | 80 |
| 4. Superior girls | 67 | 77 | 49 | 82 |
| 5. Average girls | 47 | 100 | 91 | 72 |
| 6. Slow girls | 25 | 30 | 14 | 49 |
| Totals | 293 | 455 | 329 | 425 |

[a] These N's vary for different parts of the analysis.

formula devised for determining this accuracy of perceived status score was:

$APS = CG \frac{G \times T}{W - 1} + 5$; in which:

APS = accuracy of perceived status
CG = number of correct guesses made by student
G = total number of guesses made by student
T = total number of times student was chosen by others
N-1 = number of pupils in the classroom minus himself
5 = constant added to the results to eliminate negative scores

If a student were chosen by a large number of students, there would be more probability that his guesses of who would choose him would be accurate simply by chance, than the guesses of a student who was chosen by only a few classmates. To correct for this chance occurance, the above described variation of the correction-for-guessing formula (score = $R - \frac{W}{N-1}$) was developed. In arriving at an accuracy of perceived status score this formula considered variations in the number of guesses made by the student, the number of correct guesses, the number of the times the subject was chosen by others, and the differences in class size. To eliminate negative accuracy scores, a constant five was added.

Scores on sociometric status and accuracy of sociometric perception were obtained in this manner for each of the three years. Data which were utilized in the correlational study relative to mental ability, academic achievement, attitudes, and concept of self were gathered during the third year of the study while the subjects were in the sixth grade.

## Method of Analysis

The study was concerned primarily with the effects of the two grouping treatments, ability level, and sex on sociometric status, accuracy of perceived status, and stability of sociometric status. Comparisons were made between pupils in the same grouping system and at the same level of ability, but of different sex. These comparisons permitted an appraisal of sex differences on the sociometric variables of pupils of comparable ability in the same grouping treatment. To appraise the influence of ability level, a series of comparisons were made within each grouping system by comparing pupils of the same sex, but of different levels of ability. These comparisons, in effect, held grouping method and sex constant in order to investigate the possible interactions between ability level and the sociometric variables. To determine the effects of the two grouping treatments on the sociometric variables, comparisons were made between pupils of the same sex and ability level in random grouped versus ability grouped classes.

The various combinations of the three variables of grouping treatment, ability level, and sex of pupils provide 12 groups that could be compared on each dependent variable. These comparisons were tested for significance using either single classification analysis of variance or chi square contingency tables.

4

## RESULTS

### Sociometric Choice Patterns

Pupils in the various subgroups employed in this study were classified into three sociometric choice categories: stars, regulars, and neglectee-isolates. Sociometric choice patterns were then compared using chi-square contingency tables to determine whether significant differences were present between the patterns of comparable pupils in the two treatments, between pupils of different ability within each treatment, and between comparable boys and girls in each treatment.

#### Between-treatment Differences

It will be recalled that two possible effects of ability grouping on sociometric choice patterns were discussed in the working hypotheses. One of these was that ability grouping would create a leadership vacuum in ability-grouped sections made up of average and slow pupils, thus reducing the over-all proportion of stars in District A. In order to check this possibility, the over-all distribution of stars, regulars, and neglectee-isolates for District A and R were compared. It will be noted in Table 7-4, row . that of the four over-all comparisons made, only one reached statistical significance. This was the difference between Sample IV pupils in District A and R during the first year of study. Examination of this contingency table reveals that the significant difference did in fact reflect a smaller number of stars and a greater number of regular pupils in District A than would occur by chance. These results suggest that during the first year of an ability grouping situation, some leadership vacuum was present in regular and slow groups. A similar trend in the overall sociometric choice pattern also occurred during the first year of the study for Sample VI. The trend, however, did not reach statistical significance for this sample and suggests that sixth grade pupils adapt more rapidly to the ability grouped situation than do fourth grade pupils.

No such trend, however, existed for Sample IV pupils during the second and third years of the study. Overall comparisons for these years indicate a slightly larger number of stars in District A than in District R. These data would suggest that once the child has adapted to his new role in the ability grouped class, the so-called leadership vacuum quickly disappears, and pupils of average and low ability emerge as stars.

In addition to these overall between-treatment comparisons, the sociometric choice patterns of boys and girls at each ability level in District A were compared to those of their counterparts in District R. The results of these comparisons for the three years of the study are shown in Table 7-4, rows 2-7. The reader will recall that our working hypothesis predicted a smaller number of stars among superior pupils in District A. This trend was, in fact, evident in the eight comparisons between superior pupils in the two districts. It will be noted that three of these differences reached statistical significance and one additional approached significance. The three non-significant differences also followed the same trend of fewer stars emerging among superior pupils in District A. In an ability grouping situation, superior pupils, being placed together, must compete among themselves for a smaller number of positions of social leadership than would be available to them in random grouped classrooms. It is noteworthy that, although ability grouping reduced the superior

Table 7-4

Differences Between Different Sub-Groups in the Four Replications Based Upon Chi-Square Contingency Tables

| Groups (group having more desirable distribution of sociometric classifications listed first when significant differences are present.) | Purpose of Comparison | 1958-59 4th Grade (Sample IV) | 1959-60 5th Grade (Sample IV) | 1960-61 6th Grade (Sample IV) | 1958-59 6th Grade (Sample VI) |
|---|---|---|---|---|---|
| Between-Treatment Differences | Overall Differences: To check overall proportion of sociometric classification between groups. | | | | |
| 1. All subjects District A vs. all subjects District R. | | 6.59[a] | N.S. | N.S. | N.S. |
| 2. Superior Boys, R. vs. A. | These 6 comparisons are between comparable sub-groups in terms of sex and ability in District A versus District R. | N.S. | N.S. | N.S. | 6.52 |
| 3. Superior Girls, R. Vs. A. | | 8.56 | A.S. | N.S. | 6.04 |
| 4. Average Boys, A. vs. R. | | A.S. | 8.53 | N.S. | A.S. |
| 5. Average Girls, A. vs. R. | | N.S. | N.S. | A.S. | N.S. |
| 6. Slow Boys, A. vs. R. | | N.S. | 12.87 | A.S. | N.S. |
| 7. Slow Girls, A. vs. R. | | N.S. | 9.25 | N.S. | 7.38 |
| Level Differences | | | | | |
| 8. District A: Superior vs. Average Boys | These 6 comparisons are between sub-groups of the same sex but different ability level in District A. | A.S. | N.S. | N.S. | N.S. |
| 9. District A: Superior vs. Slow Boys | | 7.25 | N.S. | N.S. | N.S. |
| 10. District A: Average vs. Slow Boys | | N.S. | N.S. | N.S. | N.S. |
| 11. District A: Superior vs. Average Girls | | N.S. | N.S. | N.S. | N.S. |
| 12. District A: Superior vs. Slow Girls | | N.S. | N.S. | N.S. | N.S. |
| 13. District A: Average vs. Slow Girls | | N.S. | N.S. | N.S. | N.S. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 14. | District R: Superior vs. Average Boys | These 6 comparisons are between sub-groups of the same sex but different ability level in District R. | N.S. | 16.37 | 12.68 | 30.09 |
| 15. | District R: Superior vs. Slow Boys | | 10.00 | 27.62 | 24.84 | 23.44 |
| 16. | District R: Average vs. Slow Boys | | N.S. | 25.12 | A.S. | N.S. |
| 17. | District R: Superior vs. Average Girls | | 12.10 | 12.93 | 29.73 | A.S. |
| 18. | District R: Superior vs. Slow Girls | | 16.11 | 21.99 | 26.55 | 27.52 |
| 19. | District R: Average vs. Slow Girls | | N.S. | N.S. | N.S. | 10.67 |
| Sex Differences | | | | | | |
| 20. | District A: Superior Boys vs. Superior Girls. | These 6 comparisons are between sub-groups of boys and girls of comparable ability in the two districts and are aimed at discovering sex differences in sociometric status. | N.S. | N.S. | N.S. | N.S. |
| 21. | District A: Average Boys vs. Average Girls. | | N.S. | N.S. | N.S. | N.S. |
| 22. | District A: Slow Boys vs. Slow Girls. | | N.S. | N.S. | N.S. | N.S. |
| 23. | District R: Superior Boys vs. Superior Girls. | | 7.32 | 10.60 | N.S. | A.S. |
| 24. | District R: Average Boys vs. Average Girls. | | N.S. | N.S. | N.S. | A.S. |
| 25. | District R: Slow Boys vs. Slow Girls | | N.S. | A.S. | N.S. | N.S. |

---

[a] Chi-square of 5.99 significant at .05 level; 9.21 significant at .01 level; 13.8 significant at .001 level.

pupil's chances of reaching the star classification, it did not increase the liklihood of his being classified as a neglectee-isolate.

It was hypothesized that larger numbers of stars would be present among average pupils in ability grouped classrooms because of the positions of social leadership opened to these pupils by removal of the superior pupils from the classroom. It will be noted in Table 7-4 that one of the eight between-treatment differences for average pupils was statistically significant and two others approached significance. These differences plus the remaining four non-significant differences were all in the hypothesized direction. That is, the contingency tables revealed a larger proportion of stars among average pupils in District A than in District R.

It was also hypothesized that a greater number of stars would emerge among slow pupils in District A because of the removal of superior and average pupils who might be expected to dominate the social situation in a heterogeneous classroom. It will be noted in Table 7-4 that three of the eight comparisons between slow pupils in the two districts were statistically significant while a fourth approached significance. Examination of the eight chi-square contingency tables for slow pupils revealed that significant differences in all cases were due to a greater number of stars and a smaller number of neglectee-isolates in District A as compared with District R. Contingency tables not yielding statistically significant chi-squares were also examined and it was found that the above trend also occurred in these tables. It is noteworthy that between-treatment differences between Sample IV slow pupils during the first year were not significant for either boys or girls, again suggesting that slow pupils require a certain amount of time for adjustment before they emerge as leaders in the ability grouped situation.

The total pattern revealed in the between-treatment comparisons is strikingly consistent. This pattern seems to indicate that average and slow pupils generally have a far better opportunity to attain positions of social leadership and are less likely to be classified as neglectee-isolates in the ability grouped situation than in a random grouped situation. Superior pupils generally have less chance of emerging as stars in the ability grouped classroom. Examination of the contingency tables indicate, however, that superior pupils do not more frequently become neglectee-isolates in the ability grouped classroom.

## Level Differences

It was hypothesized that no significant differences between District A pupils at different ability levels would be found. This hypothesis was based on the assumption that in the ability grouped situation pupils have about an equal chance of attaining favorable sociometric positions regardless of the ability level of the group in which they have been placed. It was thought that some differences between superior and slow pupils might emerge in District A during the first year of the study because of the possibility that slow pupils would be unable to exploit immediately the better opportunities for social status available to them in ability grouped classes. It will be noted in Table 7-4, rows 8-13, that of the 24 comparisons between pupils of different ability levels in District A, only one reached statistical significance at the .05 level and one approached significance. The significant difference was between the sociometric patterns of superior and slow boys and reflected more favorable patterns for superior boys. This difference again suggests that slow pupils are not capable of immediately filling the leadership vacuum that is created when ability grouped classes are established. These data also seem to indicate in a remarkably consistent manner that in the ability grouped situation pupils have an equal opportunity to attain a favored sociometric position regardless of their ability level.

It was hypothesized that significant differences would occur between pupils of different ability levels in District R. This hypothesis was based largely on previous research indicating that in the heterogeneous class the superior pupil has a greater chance of gaining peer group recognition than his less capable contemporaries. It was also hypothesized that in comparisons between average and slow pupils, average pupils would have the more favorable sociometric choice pattern but that these differences would generally be smaller than those involving comparisons with superior pupils. As superior pupils in the heterogeneous classrooms tend to monopolize positions of social leadership, the lack of opportunity for both average and slow pupils to emerge as stars would reduce the differences between these two groups.

In Table 7-4, rows 14-19 summarize the chi-square comparisons between District R pupils of different ability levels. It will be noted that the eight comparisons between superior and slow pupils were all statistically significant beyond the .01 level with seven of the eight significant beyond the .001 level. Of the eight comparisons between superior and average pupils in District R, three were significant beyond the .001 level, three were significant beyond the .01 level, one approached significance, and one was non-significant. The two non-significant differences were consistent in direction with the significant differences, favoring the superior pupils. These data seem to present overwhelming evidence that the superior pupil is in a highly favored sociometric position in the random grouped classroom as compared with average or slow pupils.

Of the eight comparisons between average and slow pupils, two were statistically significant, one approached statistical significance and the remaining five were non-significant. Both the significant and non-significant differences generally favored average pupils. These data reflect a much larger gap between superior pupils on one hand and average and slow pupils on the other than is found between the average and slow groups in terms of their success in obtaining favorable sociometric classifications in the random grouped classroom.

### Sex Differences

Comparisons between the distributions of stars, regulars, and neglectee-isolates for boys and girls of comparable ability in each district may be found in rows 20-25 of Table 7-4. It was hypothesized that no significant sex differences would emerge. It will be noted that for District A, none of the 12 between-sex comparisons yielded significant chi-square values. For District R, two of the between-sex comparisons reached statistical significance. Both of these were between superior boys and girls, and in both cases, the girls obtained the more favorable distribution of sociometric choices. Three sex differences in District R approached significance, one favoring girls and the remaining two favoring boys.

## Accuracy of Perceived Status

### Between-treatment Differences

The reader will recall that accuracy of perceived status is a measure of agreement between the pupils' prediction of who would choose him as a "best friend" and the actual "best friend" choices he received. The mean scores of the different groups involved in the study were compared using analysis of variance.

The mean accuracy of perceived status scores for the 12 groups compared in this phase of the research are found in Table 7-5. A total of 18 comparisons were made between pupils of comparable sex and ability in District A and District R. These between-treatment differences are summarized in Table 7-6. It will be noted that of the 18 comparisons, three reached statistical significance. Two of these differences were between average boys in the two districts. During the second year of this study, District R average boys scored significantly higher on this variable, but in the final year, comparison of the same pupils showed the District A group to be significantly higher. Comparisons between slow boys showed the District A sample to be significantly higher during the second year of study. Differences for the other two years, though not significant, favored District A during the first year and District R during the final year. Review of Table 7-5 for between-treatment trends indicates that the non-significant differences did not consistently favor either district. These data would suggest that the differences found were either due to chance or represent some unique characteristic of the samples concerned. It may be concluded that no significant, meaningful trend in the between-treatment differences is evident in accuracy of perceived status scores.

### Level Differences

A total of 36 within-treatment comparisons between pupils at different ability levels are summarized in Table 7-7. Of the 36 differences between group means, 17 reached statistical significance with nine of these being significant beyond the .01 level. All 17 of these significant differences were in the direction of higher mean accuracy scores for pupils of higher ability. It may be noted in Table 7-5 that 16 of the remaining 19 non-significant differences were also consistent with this trend, with higher mean accuracy scores being obtained by pupils of higher ability. The significant within-treatment differences occurred somewhat more frequently in District R than District A but both districts showed a consistent pattern. It is also noteworthy that the number of significant differences increased during the three years of the study, suggesting that brighter students gained more rapidly in their ability to perceive accurately their sociometric status than did pupils of lesser ability.

### Sex Differences

Only one of the 18 differences between the mean accuracy of perception scores of boys and girls within each district reached statistical significance. Comparisons between the scores of boys' and girls' groups in Table 7-5 reveals that comparable groups usually obtained very similar mean scores with no consistent trend for either boys or girls to be higher. These data would suggest that no sex differences occurred in accuracy of perceived status during the intermediate grades of elementary school.

### Table 7-5

### Subgroup Mean Scores for Accuracy of Perceived Status

| Group | Ns and Mean Scores | | | | | |
|---|---|---|---|---|---|---|
| | 1958-1959 | | 1959-1960 | | 1960-1961 | |
| | N | M | N | M | N | M |
| District R | | | | | | |
| 1. Superior Boys | 121 | 6.730 | 121 | 6.817 | 165 | 6.914 |
| 2. Average Boys | 79 | 6.577 | 110 | 6.902 | 116 | 6.78 |
| 3. Slow Boys | 59 | 6.411 | 80 | 5.958 | 91 | 6.56 |
| 4. Superior Girls | 136 | 6.759 | 136 | 6.912 | 163 | 7.07 |
| 5. Average Girls | 72 | 6.433 | 113 | 6.66 | 122 | 6.78 |
| 6. Slow Girls | 29 | 6.228 | 35 | 6.219 | 45 | 6.38 |
| District A | | | | | | |
| 7. Superior Boys | 45 | 6.889 | 60 | 6.941 | 50 | 6.93 |
| 8. Average Boys | 56 | 6.305 | 129 | 6.571 | 88 | 7.34 |
| 9. Slow Boys | 53 | 6.258 | 59 | 6.382 | 37 | 6.19 |
| 10. Superior Girls | 67 | 6.751 | 77 | 6.636 | 49 | 7.12 |
| 11. Average Girls | 47 | 6.643 | 100 | 6.688 | 91 | 7.02 |
| 12. Slow Girls | 25 | 6.416 | 30 | 6.492 | 14 | 6.71 |

Table 7-6

Between-treatment Differences in Accuracy of Perceived Status

| Groups | 1958-1959 | Significant t-tests<br>1959-1960 | 1960-1961 |
|---|---|---|---|
| Superior boys | N.S.[a] | N.S. | N.S. |
| Average boys | N.S. | R-2.36[b] | A-3.68 |
| Slow boys | N.S. | A-2.29 | N.S. |
| Superior girls | N.S. | N.S. | N.S. |
| Average girls | N.S. | N.S. | N.S. |
| Slow girls | N.S. | N.S. | N.S. |

[a] Difference not significant (N.S.)

[b] Indicates that District R average boys earned a significantly higher mean accuracy of perceived status score than District A average boys during the 1959-1960 school year, with a t of 2.36.

It seems safe to assume that the one significant difference obtained for each district was due to chance.

## Mobility in Sociometric Status

In order to determine mobility of sociometric choice status over the three year period of study, the sociometric classification (star, regular, or neglectee-isolate) was determined for each pupil during each of the three years of the study. Pupils were then classified into three groups: those moving to a higher sociometric category, those moving to a lower sociometric category, and those staying the same from one year to the next. Comparisons were made between the first and second years, second and third years, and first and third years. After mobility data had been obtained for each subgroup, comparisons between the patterns of change in sociometric classification between different subgroups were made. Chi-square contingency tables were used to determine the direction and level of significance of between-treatment, within-treatment, and sex differences in sociometric choice mobility. When significant differences in patterns of changes were obtained, the chi-square contingency tables were examined to determine the nature of these differences.

### Between-treatment Differences

Between-treatment comparisons first were made between the total samples in the two districts. Three chi-square contingency tables comparing the total samples in Districts A and R on sociometric mobility from the first to second year, second to third year, and first to third year were analyzed. The difference in sociometric status mobility between Districts A and R from the first to the second year was statistically significant beyond the .001 level (chi-square of 14.94). Differences in mobility patterns between years two and three and years one and three were not statistically significant. Examination of the significant contingency table indicated that District A pupils were significantly more mobile during this period. This mobility was in an upward direction for the most part and reflects the marked readjustments in sociometric status that occur when an ability grouping system is adopted. This same pattern, though much less pronounced, was also evident in the contingency table for years one to three. Differences between the second and third year were much smaller. This suggests that by the second year, adjustments in sociometric status that were brought about as a result of introducing the ability grouping system had largely taken place and the new status pattern that had emerged might be expected to remain relatively stable.

In order to get additional insight into between-treatment differences in sociometric mobility, these data were analyzed further by comparing the sociometric status patterns of pupils of the same ability level under the two grouping systems using chi-square contingency tables. Because of a general absence of sex differences on this variable, boys and girls were combined for this analysis. The results of the chi-square comparisons may be found in Table 7-8. It will be noted that in addition to comparing superior, average, and slow pupils in the two districts, pupils in each district who moved downward or upward in ability level were also compared to determine whether such movement in ability classification was accompanied by changes in sociometric status.

Comparison of pupils of superior ability in the two districts yielded no significant differences in mobility of sociometric status. There was a slight trend evident

Table 7-7

Significant Differences Between Mean Scores on Accuracy of Perceived Status for Pupils of the Different Levels of Ability

| Groups | Significant t-tests[a]<br>1958-1959 | <br>1959-1960 | <br>1960-1961 |
|---|---|---|---|
| District R: | | | |
| Superior vs. Average boys | N.S. | N.S. | N.S. |
| Superior vs. Slow boys | N.S. | 5.51 | 2.53 |
| Average vs. Slow boys | N.S. | 5.94 | N.S. |
| Superior vs. Average girls | 2.08 | N.S. | 2.27 |
| Superior vs. Slow girls | 2.40 | 3.36 | 3.83 |
| Average vs. Slow girls | N.S. | 2.11 | 2.13 |
| District A: | | | |
| Superior vs. Average boys | 2.72 | 2.18 | 2.15 |
| Superior vs. Slow boys | 2.91 | 2.82 | 3.15 |
| Average vs. Slow boys | N.S. | N.S. | 5.42 |
| Superior vs. Average girls | N.S. | N.S. | N.S. |
| Superior vs. Slow girls | N.S. | N.S. | N.S. |
| Average vs. Slow girls | N.S. | N.S. | N.S. |

[a] All significant differences shown on this table favor the group of higher ability level.

Table 7-8

Between-treatment Differences in Sociometric Status Mobility at Each Ability Level

| Group | Significant Chi Square Values<br>Year 1 - 2<br>1958-59<br>1959-60<br>t-value | Year 2 - 3<br>1959-60<br>1960-61<br>t-value | Year 1 - 3<br>1958-59<br>1960-61<br>t-value |
|---|---|---|---|
| Superior pupils | N.S. | N.S. | N.S. |
| Average pupils | A[a]-10.97[b] | N.S. | N.S. |
| Slow pupils | A-12.39[b] | N.S. | A-10.48[b] |
| Pupils moving downward in ability level | A-11.67[b] | N.S. | N.S. |
| Pupils moving upward in ability level | R-6.46[c] | N.S. | R-10.43[b] |

[a] Indicates that District A sample made a greater proportion of favorable changes.
[b] Significant at .01 level of confidence.
[c] Significant at .05 level of confidence.

between years one and two for District A superior pupils to make more status gains and less status losses than were made by District R superior pupils during this same period. This may reflect a tendency for superior pupils to regain some of the sociometric status they initially lose in an ability grouped situation.

Comparisons between the sociometric status mobility patterns of average pupils show District A pupils to be significantly more mobile between years one and two than comparable District R pupils. This mobility was generally upward and reflects an improvement of the average pupils' status during the second year of their experience in the ability group situation. These data suggest that the introduction of ability grouping provided an opportunity for many pupils of average ability to assume new and heretofore unobtainable leadership roles in the social structure of the classroom.

The difference between pupils of low ability in the two districts in mobility of sociometric status reached statistical significance in comparisons between years one and two and years one and three. Examination of the chi-square contingency tables for slow pupils reveals a consistent tendency for pupils in ability grouped classrooms to experience more upward mobility of sociometric status from one year to the next than comparable random grouped pupils. These data seem to support the conclusion that the slow pupil has a greater opportunity for gaining social approval and recognition in an ability grouped classroom than he does in a random grouped classroom. The absence of any significant between-treatment differences in mobility between the second and third year of the study again suggests that adjustments in sociometric status brought about by ability grouping are largely completed by the second year after its introduction.

In the chi-square analysis of pupils who received a lower ability level classification in a succeeding year, it will be noted that a significant difference was obtained between the first and second years. Examination of the chi-square contingency table reveals that District A pupils in this classification experienced significantly more upward mobility and less downward mobility than was experienced by the comparable District R pupils. This same trend prevailed in comparisons between the second and third and first and third years although neither attained statistical significance. The movement downward from one ability level classification to another in the ability grouped situation would usually result in a pupil being at the upward extreme of his new ability classification. Thus, the downward movement seems to increase his relative status as he is recognized as having come from a higher ability group and is near the top of his new group whereas he was near the bottom of his previous group. In randomly grouped classrooms it will be recalled that the pupil's ability level classification was unknown to him. He was not subject to room reassignment as a result of obtaining a lower classification. These data would suggest that any damage to the ability grouped pupil's self-concept that might occur because of his reclassification to a lower ability section might be compensated to some extent by gains in status he is likely to experience in his new class.

Finally, pupils in District A and District R who moved upward in ability level classification were studied. Two of the three chi-square comparisons made between these pupils yielded statistically significant differences. In each instance, pupils in random grouped classrooms gained in sociometric acceptance to a greater extent and lost status to a lesser degree than comparable pupils in District A classrooms. This difference may well have been caused by the same factor which appeared to operate on ability grouped pupils who moved downward in ability level. That is, when the pupil is moved from the upper extreme of one ability group and placed in the lower extreme of the next higher group his relative status within the classroom may suffer. These findings suggest that although ability grouping might initially provide greater opportunities for social leadership for those of average and low ability, it might also make more difficult the task of maintaining or gaining social recognition if the pupil makes achievement gains sufficient to raise his ability level classification. Loss of social status that is likely to accompany his reassignment to a new classroom might outweigh any gain in status he may experience from being identified as a member of the brighter group.

In reviewing the data given in Table 7-8, it will be noted that six of the 15 between-treatment comparisons attained statistical significance. Most of the non-significant comparisons revealed a trend consistent with the direction of the significant findings (see Table 7-9). These differences were in the direction of greater upward mobility of sociometric status from one year to the next for pupils in ability group classrooms and greater stability of status for pupils in random grouped classrooms. On the basis of these findings, it is necessary to conclude that a real difference did exist between the two districts in stability of sociometric status.

### Level Differences

A total of 18 comparisons were made between pupils of different ability levels within the two grouping treatments. Only one of these comparisons attained statistical significance at the .05 level suggesting that no true within-treatment differences in stability of sociometric status existed.

### Sex Differences

A total of 30 comparisons were made between the stability of sociometric status of boys and girls of comparable ability. Only two significant differences were obtained in these comparisons, with District R average boys showing significantly greater stability in one comparison and District A superior boys showing greater stability in the other significant difference. Because only two statistically significant differences were obtained out of a total of 30 comparisons and because the non-significant differences did not follow any consistent trend, it appears probable that the two significant sex differences obtained were due to chance. Thus, it may be concluded that no true sex differences were present among pupils in the intermediate grades in stability of sociometric status.

### Long Range Stability of the Accuracy of Perceived Status Score

In order to determine the long range stability of the accuracy of perceived status scores, accuracy scores on each of the three sociometric criteria were correlated between the first and second year, second and third year, first and third year. The correlation coefficients obtained were generally low, although 11 of the 18 were statistically significant, ranging from .13 to .29. These correlations are presented in Table 7-10. It will be noted that all six of the correlations that were based on the first sociometric criterion, best friend nominations, were significant.

In view of the low but consistently significant long range stability of the first sociometric criterion, further analysis was carried out with this criterion for pupils at each level of ability. These data are summarized in Table 7-11. It will be noted that seven of the 18 correlations reached statistical significance, ranging from .30

Table 7-9

Mobility and Stability of Sociometric Status

| | Percentage of Total Group in Each Category | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Upward Mobile | | | | Stable | | | | Downward Mobile | | | |
| Group | Neg. to Reg. | Neg. to Star | Reg. to Star | Total % Up | Star to Star | Reg. to Reg. | Neg. to Neg. | Total % Stable | Star to Reg. | Star to Neg. | Reg. to Neg. | Total % Down |
| **District R** | | | | | | | | | | | | |
| Superior | | | | | | | | | | | | |
| Yr. 1-Yr. 2 | 07 | 00 | 08 | 15 | 21 | 28 | 07 | 56 | 16 | 02 | 11 | 29 |
| Yr. 2-Yr. 3 | 10 | 02 | 13 | 25 | 20 | 27 | 11 | 58 | 10 | 00 | 07 | 17 |
| Average | | | | | | | | | | | | |
| Yr. 1-Yr. 2 | 07 | 01 | 09 | 17 | 07 | 32 | 30 | 69 | 07 | 00 | 07 | 14 |
| Yr. 2-Yr. 3 | 15 | 01 | 08 | 24 | 05 | 28 | 23 | 56 | 03 | 01 | 15 | 19 |
| Slow | | | | | | | | | | | | |
| Yr. 1-Yr. 2 | 09 | 00 | 00 | 09 | 02 | 18 | 48 | 68 | 05 | 00 | 18 | 23 |
| Yr. 2-Yr. 3 | 23 | 01 | 03 | 27 | 04 | 16 | 40 | 60 | 01 | 00 | 12 | 13 |
| **District A** | | | | | | | | | | | | |
| Superior | | | | | | | | | | | | |
| Yr. 1-Yr. 2 | 12 | 02 | 07 | 21 | 15 | 27 | 19 | 61 | 12 | 00 | 06 | 18 |
| Yr. 2-Yr. 3 | 15 | 02 | 14 | 31 | 15 | 21 | 21 | 57 | 03 | 03 | 06 | 12 |
| Average | | | | | | | | | | | | |
| Yr. 1-Yr. 2 | 25 | 04 | 13 | 42 | 09 | 11 | 19 | 39 | 06 | 02 | 11 | 19 |
| Yr. 2-Yr. 3 | 11 | 02 | 11 | 24 | 08 | 34 | 21 | 63 | 10 | 00 | 03 | 13 |
| Slow | | | | | | | | | | | | |
| Yr. 1-Yr. 2 | 18 | 13 | 11 | 42 | 05 | 24 | 18 | 47 | 00 | 00 | 11 | 11 |
| Yr. 2-Yr. 3 | 13 | 00 | 11 | 24 | 10 | 29 | 18 | 57 | 08 | 00 | 11 | 19 |

to .48. This analysis was carried out primarily to give the reader further insights into the nature of the accuracy of perceived status score. Measures related to this variable have been employed to a very limited degree in previous research. The results suggest that this characteristic, as measured with the tools in this project, was consistent at a low but statistically significant level over the three year period.

## Relationships Between Sociometric Status and Other Variables

Correlations between sociometric status and accuracy of perceived status were computed for both districts. The same correlation, .52, was obtained for both District R (N = 708) and District A (N = 329). This correlation indicates a moderate relationship between the variables and is significant beyond the .01 level of confidence. The reader will recall that sociometric status was based on nominations of others while accuracy of perceived status was based on the pupil's estimate of who would choose him as best friend. It appears unlikely that these two measures are tapping the same variable, but it may be that accuracy of sociometric perception tends to contribute to the individual's sociometric status. In other words, insight into the attitudes of others could assist the pupil in developing friendships. A number of other studies have found positive correlations between social sensitivity and social acceptance (Gronlund, 1955; Dymond, Hughes and Raabe, 1952; Ausubel and Schiff, 1955; Williams, 1962' Ausubel, Schiff and Gasser, 1952).

Correlations between sociometric status and mental ability measured by California Test of Mental Maturity showed significant relationships for District R pupils on both the language and the non-language scores, while relationships between these variables in District A were not significant. These findings show that mental ability is related to sociometric status in random grouped classes but no relationship exists between these variables in ability group classrooms. These results suggest that the relationships between mental ability and sociometric status reported in previous research may be dependent upon the type of classroom organization. A number of researchers have reported that children with higher levels of intellectual ability tend to receive more sociometric choices than children of lower ability (Gallagher, 1958; Gronlund, 1959; Laughlin, 1954; Bonney, 1942, 1943, 1944). It is, of course, difficult to interpret relationships between mental ability and sociometric choice in ability grouped classrooms because pupils in such classrooms are limited in their choice to other pupils in a restricted ability range.

Correlations were computed between sociometric status and academic achievement, using a composite score based on four of the Sequential Tests of Educational Progress. The results followed the same pattern as the correlations with mental ability, with a significant relationship found in the District R sample and no significant relationship for the District A Sample. Correlations between sociometric status and both intelligence and achievement are probably reduced by the restriction of ability range in the District A classrooms. These data may suggest, however, that in ability grouped classrooms, the pupil's ability or achievement level is not a factor in his sociometric status, while these variables appear to exert some influence on status in random grouped classrooms.

Correlations between sociometric status and the attitudes of pupils toward school, their teacher, and peers were also computed. It will be noted in Table 7-12 that low but significant relationships generally existed between desirable attitudes and sociometric status with the highest correlations being found between sociometric status and attitudes towards peers. As the pupil's attitude towards his peers is almost

Table 7-10

Correlation Coefficients Between Accuracy of Perceived Status Scores Obtained in Different Years

| Group | Correlation Coefficients<br>1st Criterion | <br>2nd Criterion | <br>3rd Criterion |
|---|---|---|---|
| Year 1 - Year 2 | | | |
| Dist. R (N=389) | .21[a] | .07 | .00 |
| Dist. A (N=224) | .29[a] | .20[a] | .21[a] |
| Year 2 - Year 3 | | | |
| Dist. R (N=529) | .28[a] | .08 | .04 |
| Dist. A (N=294) | .24[a] | .06 | .04 |
| Year 1 - Year 3 | | | |
| Dist. R (N=357) | .17[a] | .13[a] | .06 |
| Dist. A (N=131) | .20[b] | .22[a] | .24[a] |

[a] Significant at the .01 level of confidence.
[b] Significant at the .05 level of confidence.

Table 7-11

Test-Retest Correlation Coefficients of First Criterion Accuracy of Perceived Status Scores

| | Superior | | Average | | Slow | |
|---|---|---|---|---|---|---|
| Group | N | r | N | r | N | r |
| Year 1 - Year 2 | | | | | | |
| Dist. R. | 123 | .11 | 68 | .30[b] | 45 | .04 |
| Dist. A. | 67 | .31[a] | 53 | .12 | 38 | .43[a] |
| Year 2 - Year 3 | | | | | | |
| Dist. R. | 207 | .35[a] | 140 | .00 | 75 | .48[a] |
| Dist. A. | 66 | .12 | 124 | .31[a] | 38 | .07 |
| Year 1 - Year 3 | | | | | | |
| Dist. R. | 120 | .15 | 51 | .11 | 32 | .07 |
| Dist. A. | 34 | .37[b] | 28 | .00 | 23 | .30 |

[a] Significant at the .01 level of confidence.
[b] Significant at the .05 level of confidence.

certainly influenced by his sociometric status, it is not surprising that a significant correlation was found between these two variables.

Low but significant relations were also found between the pupil's sociometric status and the various aspects of self concept as measured by Bills' Index of Adjustment and Values. It will be noted, however, that these correlations are very low, reflecting very slight relationships between the measures. The presence of these relationships, however, suggests that somewhat larger correlations may exist between the characteristics themselves. This speculation, however, must await the development of more sophisticated measures before it may be tested.

## SUMMARY AND CONCLUSIONS

The sociometric data collected for Sample IV pupils during their fourth, fifth, and sixth grade years, and for Sample VI pupils during their sixth grade year seems to provide considerable evidence in support of the following conclusions relating to sociometric choice in ability grouped and random grouped classes during the intermediate grades:

1. The overall proportions of stars, regulars, and neglectees in District A and District R were similar, thus leading to the conclusion that ability grouping does not result in a permanent leadership vacuum in groups of average and slow pupils. Evidence suggests, however, that some time is required for ability grouped average and slow fourth grade pupils to emerge as sociometric stars, thus creating a temporary social leadership vacuum at this grade level. No such temporary vacuum was evident for Sample VI during the first year of the study, suggesting that this temporary leadership vacuum if present at all, is of much shorter duration for sixth grade pupils.
2. Superior students lose some sociometric status when placed in ability-grouped classrooms. This loss is particularly evident in the star classification but is not accompanied by any increase in the neglectee-isolate classification. Data on mobility of sociometric status suggests that superior pupils regain some of this status loss after they have made an adjustment to the ability grouped situation.
3. Average and slow pupils appear to have a far better chance of gaining social recognition in ability grouped classrooms than do comparable pupils in random grouped classrooms. For the slow pupil, ability grouping not only appears to increase the pupil's chances of being classified as a star, but also reduces his chances of being classified as a neglectee-isolate.
4. In random grouped classrooms, the pupil's ability level appears to be an important factor in determining his sociometric status. Sociometric status patterns differ considerably for pupils of different ability levels in random grouped classrooms, with pupils of higher ability consistently obtaining more favorable mean scores.
5. Ability does not appear to be an important factor in determining sociometric status in ability grouped classrooms. In the Utah study sociometric status patterns for pupils of different ability levels in ability grouped classrooms were closely comparable, showing no evidence of real difference between the mean sociometric choice scores of pupils at different ability levels.
6. The presence of significant correlations obtained when sociometric status was related to ability and achievement for District R pupils further demonstrates the importance of ability to status in the random grouped class. The absence of such correlations suggests that these factors are far less important to friendship choices in ability grouped classrooms.
7. Sociometric status mobility patterns between District A and R demonstrate a

Table 7-12

Correlations Between Sociometric Status and Other Variables for Each District

| | Total Sociometric Status | |
|---|---|---|
| Other Variables | District R | District A |
| Mental Ability | | |
| Language Mental Age | .396[a] | .052 |
| Non-Language Mental Age | .239[a] | .002 |
| Academic Achievement | .161[a] | .085 |
| Desirable Attitudes | | |
| Toward School | .234[a] | .121[b] |
| Toward Teachers | .099[b] | .113 |
| Toward Peers | .355[a] | .311[a] |
| Concept of Self | .161[a] | .194[a] |
| Acceptance of Self | .084[b] | .141[b] |
| Ideal Self | .226[a] | .129[b] |
| Discrepancy | .013 | .079 |

[a] Significant at .01 level of confidence.
[b] Significant at .05 level of confidence.

greater number of sociometric changes in ability grouped classrooms. The trend of these changes is generally towards higher sociometric status, leading to the conclusion that organization of pupils into ability grouped classes brings about marked changes in sociometric status patterns during the first one or two years and generally leads to permanently higher mean sociometric status for average and slow groups.

8. The differences between ability grouping and random grouping did not appear to have any influence upon the pupil's ability to accurately perceive his sociometric status, i.e., to accurately predict which pupils in his classroom group would choose him as a best friend.

9. The pupil's ability to perceive accurately his sociometric acceptance by others is related to academic ability in both districts. Pupils of higher academic ability probably are more perceptive of social and interpersonal relationships and this greater perception may assist them in gaining the acceptance and friendship of others.

10. The lack of any consistent significant differences between boys and girls suggests that the effects of ability grouping and random grouping on sociometric variables are about the same for both sexes. Thus, it may be concluded that sociometric status, accuracy of perceived status, and stability of sociometric status are essentially unrelated to sex in either random grouped or ability grouped classrooms at the intermediate grade levels of the elementary school.

11. Low but significant correlations exist between sociometric status, pupil attitude, and self-concept variables. These results, though low, demonstrate the presence of relationships between the pupil's social acceptance and his perception of himself, his school environment, and his peers.

## REFERENCES

Ausubel, D. P., & Schiff, H. M. Some intrapersonal and interpersonal determinants of individual differences in socioempathic ability among adolescents. J. soc. Psychol., 1955, 41, 39-56.

Ausubel, D. P., Schiff, H. M., & Gasser, E. B. A preliminary study of developmental trends in socioempathy: accuracy of perception of own and others' sociometric status. Child Develpm., 1952, 23, 111-128.

Barbe, W. B. Peer relationships of children of different intelligence levels. Sch. Soc., 1954, 80, 60-62.

Barbe, W. B. Evaluation of special classes for gifted children. Except. Children, 1955, 22, 60-62.

Bedoian, V. H. Social acceptability and social rejection of the underage, at age, and overage pupils in the sixth grade. J. educ. Res., 1954, 47, 513-520.

Bell, M. E. A comparative study of mentally gifted children heterogeneously and homogeneously grouped. Dissertation Abstr., 1959, 19, 2509.

Bettelheim, B. Segregation: new style. *Sch. R.*, 1958, 66, 251-272.

Biehler, R. F. Companion choice in the kindergarten. *Child Develpm.*, 1954, 25, 45-50.

Bjerstedt, A. The interpretation of sociometric status scores in the classroom. *Acta psychol., Amsterdam,* 1956, 12, 1-14.

Blanchard, B. E. Social acceptance studies of pupils in public schools. *J. educ. Res.*, 1947, 40, 503-512.

Bonney, M. E. A study of the relation of intelligence, family size, and sex differences with mutual friendships in the primary grades. *Child Develpm.*, 1942, 13, 79-100.

Bonney, M. E. The relative stability of social, intellectual, and academic status in grades II to IV, and the inter-relationships between these various forms of growth. *J. educ. Psychol.*, 1943, 34, 88-102.

Bonney, M. E. Relationships between social success, family size, socioeconomic home background, and intelligence among school children in grades III to V. *Sociometry*, 1944, 7, 26-39.

Bonney, M. E. A study of the sociometric process among sixth-grade children. *J. educ. Psychol.*, 1946, 37, 259-372.

Bronfenbrenner, U. The measurement of sociometric status, structure and development. *Sociometry Monogr.*, 1945, 6.

Brueckner, L. J., & Bond, G. P. *The diagnosis and treatment of learning difficulties.* New York: Appleton-Century-Crofts, 1955.

Buswell, M. M. The relationship between the social structure of the classroom and the academic success of the pupils. *J. exp. Educ.*, 1953, 22, 37-52.

Carew, D. K. A comparison of activities, social acceptance, and scholastic achievement of men-students. *Personnel guid. J.*, 1957, 36, 121-124.

Cavanaugh, M. H. Grouping in depth. *Clearing House*, 1960, 34, 427-428.

Challman, R. Factors influencing friendships among preschool children. *Child Develpm.*, 1932, 3, 146-158.

Cunningham, R. *et al*. *Understanding group behavior of boys and girls*. New York: Columbia University, Bureau of Publications, 1951.

Davis, J. A. Correlations of sociometric status among peers. *J. educ. Res.*, 1957, 50, 561-569.

Dineen, M. A., & Gary, R. Effects of sociometric seating on a classroom cleavage. *Elem. Sch. J.*, 1955, 56, 358-362.

Dymond, R. F., Hughes, A. S., & Raabe, V. L. Measureable changes in empathy with age. *J. consult. Psychol.*, 1952, 16, 202-206.

Furfey, P. H. Some factors influencing the selection of boys' chums. *J. appl. Psychol.*, 1927, 11, 47-51.

Gallagher, J. J. Peer acceptance of highly gifted children in elementary school. *Elem. Sch. J.*, 1958, 58, 465-470.

Gallagher, J. J., & Crowder, T. The adjustment of gifted children in the regular classroom. *Except. Children*, 1957, 23, 306-319.

Ginzberg, E. *Values and ideals of American youth*. New York: Columbia University, 1961.

Goldberg, M. L. Research on the talented. *Teachers Coll. Rec.*, 1958, 60, 150-163.

Goldberg, M. L. Research on the gifted. In Shertzer (Ed.) *Working with superior talented students*. Chicago: Science Research Associates, 1960.

Goldberg, M. L., Passow, A. H., & Lorge, I. Social consequences of special education for the talented. *Ed. Digest*, 1958, 24, 11-14.

Goldworth, M. The effects of an elementary school fast-learner program on children's social relationships. *Except. Children*, 1959, 26, 59-63.

Grace, H. A., & Booth, N. L. Is the gifted child a social isolate. *Peabody J. Educ.*, 1958, 35, 195-196.

Gronlund, N. E. *Sociometry in the classroom*. New York: Harper and Brothers, 1959.

Gronlund, N. E., & Whitney, A. P. Relation between pupil's social acceptability in the classroom, in the school, and in the neighborhood. *Sch. R.*, 1955, 64, 267-271.

Hamalainen, A. E. Method of grouping pupils should provide normal social situations. *Nations Sch.*, 1950, 45, 34-35.

Heber, R. F. The relation of intelligence and physical maturity to social status of children. *J. educ. Psychol.*, 1956, 47, 158-162.

Jackson, G. T. Each according to his ability. *Sch. Executive*, 1943, 62, 37-38.

Jenkins, G. G. Factors involved in children's friendships. *J. educ. Psychol.*, 1931, 22, 440-448.

Jennings, H. *Leadership and isolation*. New York: Longmans Green, 1950.

Justman, J. Personal and social adjustment of intellectually gifted accelerants and non-accelerants in junior high schools. Sch. R., 1953, 61, 468-478.

Laughlin, F. The peer status of sixth and seventh grade children. New York: Bureau of Publications, Teacher's College, Columbia University, 1954.

Lawson, D. E. An analysis of historic and philsophic considerations for homogeneous grouping. Ed. Adm. & Sup., 1957, 43, 257-270.

Luchins, A. S., & Luchins, E. H. Children's attitudes toward homogeneous groupings. Pedagog. Sem., 1948, 72, 3-9.

McGaughy, J. R. Homogeneous grouping of pupils. Childh. Ed., 1930, 6, 291-296.

McKinney, J. C. A comparison of the social psychology of G.H. Mead and J. L. Moreno. Sociometry, 1947, 10, 388-349.

Mann, H. How real are friendships of gifted and typical children in a program of partial segregation? Except. Children, 1957, 23, 199-201.

Martinson, Ruth A. The California study of programs for gifted pupils. Except Children, 1959, 36, 339-343.

Morgan, H. G. Social relationships of children in a war-boom community. J. educ. Res., 1946, 40, 271, 286.

Nelson, E. A., & Carlson, E. F. Special education for gifted children: III. Evaluation at the end of three years. J. Except. Child, 1945, 12, 6-13, 24.

Ohlsen, M. M. Guidance, an introduction. New York: Harcourt Brace, 1955.

Otto, H. J. Elementary school organization and administration. New York: Appleton-Century-Crofts, 1954.

Rochfort, G. B., Jr. Evaluating social-personal outcomes in differentiated instruction. J. Ed., 1959, 142, 66-71.

Routh, M. R. Sociometric aids. Ed., 1960, 80, 307-310.

Ryan, R. J., & Davie, J. S. Social acceptance, academic achievement, and academic aptitude among high school students. J. educ. Res., 1958, 52, 101-106.

Saville, A. A survey of ability grouping. Dissertation Abstr., 1961, 22, 1473.

Staker, A. M. Changes in social status of elementary-school pupils. Educ. Res. Bull., 1948, 27, 157-159.

Symonds, P. M., & Jackson, C. E. Measurement of the personality adjustment of high school pupils. New York: Teachers College, Columbia University, 1935.

Taylor, E. A. Some factors relating to social acceptance in eight grade classrooms. J. educ., Psychol., 1952, 43, 257-272.

Thorpe, J. G. An investigation into some correlates of sociometric status within school classes. Sociometry, 1955, 18, 49-61.

Tonsor, C. A. Must we segregate? Nat. Assn. Sec. Sch. Prin. Bul., 1953, 37, 75-77.

Williams, J. E. Acceptance by others and its relationship to acceptance of self and others: a repeat of Fey's study. J. abnorm. soc. Psychol., 1962, 65, 438-442.

Wrightstone, W. J., Justman, J., & Robbins, I. Evaluation in modern education. New York: American Book, 1956, 199-214.

## Chapter 8

# PUPIL ATTITUDES

by Walter R. Borg and Val R. Christensen[1]

## INTRODUCTION

The possible effects of ability grouping on the pupil's attitudes towards school has been a major concern of educators who have been interested in employing ability grouping in their schools. The atmosphere in which the pupil finds himself seems to differ considerably in the ability grouped and random grouped situation as well as being somewhat different at the different levels of the ability grouped situation. Thus, one might well expect to find differences in the manner that pupils react to and perceive the school environments that these treatments provide. Among the aspects of school adjustment measured in the Utah study are included measures of the pupil's school relationships, his attitudes towards the school, the teacher, and his classmates, and problems he percieves in his school experience.

## WORKING HYPOTHESES

The reader will recall that pupils in both District A and R were classified into ability levels in the Utah study on the basis of a composite achievement test battery. Thus, the bright pupil who is achieving considerably under his capacity may not be included in the superior classification. This method of grouping pupils seemed likely to have had a considerable effect upon the attitude data. High achievers in both districts are probably pupils whose value systems place considerably emphasis upon school achievement, whose psychological needs are met by achievement, and who would therefore be expected to have favorable attitudes towards school and related topics regardless of the grouping system. The fact that such pupils are achieving at a high level indicates that the school situation is filling a need for these pupils. A frequently mentioned claim of the advocates of ability grouping is that the bright pupil is frustrated in the random grouped situation. If this is the case, these frustrated pupils would be expected to have less favorable attitudes than pupils for whom the school situation provide a means of need fulfillment. However, it is possible that in the Utah study some such pupils, would not be placed in the superior classification because of the lower achievement that might be expected to accompany their frustration. Thus, their negative attitudes would show up in the average group instead of the superior group. Because the aforementioned phenomenon may have occurred, the reader should keep in mind that the working hypotheses presented were based on the assumption that the superior (high achieving) classification may not contain all superior pupils in the two districts.

---

[1] Portions of this chapter are based upon Mr. Christensen's Master's thesis which was carried out under Dr. Borg's supervision.

Three related attitude areas are sampled on the USU School Inventory. Working hypotheses relative to these three scales and the other related measures used in the Utah study will now be discussed.

### Attitude Towards Peers

Superior District R pupils might be expected to have more favorable attitudes toward peers because of the greater opportunities these pupils have of gaining social recognition in the random grouped classroom. On the other hand, however, superior pupils in District A probably have somewhat more in common with their peers such as similarity of achievement values, better identification with each other, and greater mutual respect that could lead to more favorable attitudes. It was considered likely that these two influences would offset each other thus leading to the working hypothesis that no significant differences in attitudes towards peers would be found in comparisons between superior groups in the two grouping treatments. Both of the aforementioned influences appear to operate in the same direction for average and slow pupils. That is, these pupils would have greater opportunities for peer recognition in the ability grouped classroom and also would experience more similarity between their interests and values and those of their peers of comparable ability. Thus, it was hypothesized that District A pupils of average and low ability would have more favorable attitudes than comparable District R pupils. Greater differences would be expected to occur between the slow groups, because of the relatively greater improvement of the ability grouped pupil's peer group status at this ability level.

The factors hypothesized to influence between-treatment differences in attitudes towards peers would tend to reduce differences between the attitudes of District A pupils at different ability levels. Thus, it was hypothesized that no within-treatment differences would occur between District A groups on this score. The sociometric research reported in Chapter 7, however, indicated that in random grouped classrooms peer status is related to ability and it was therefore hypothesized that significant differences favoring the groups of higher ability level would be found in the within-treatment comparisons between District R groups. No significant sex differences in attitudes towards peers were hypothesized.

### Attitude Towards the Teacher

It would be expected that attitudes toward the teacher would be more favorable in situations where the teacher can give more attention to individual pupil problems and in situations where the curriculum makes it possible for the teacher to lead the pupil to a greater number of satisfying or need-fulfilling experiences and a smaller number of frustrating experiences.

Ability grouping has been advocated because of the apparent opportunities for the teacher to provide such situations in more homogeneous classes. Thus, it was hypothesized that both superior and slow pupils would have significantly more favorable attitudes towards their teachers in ability grouped classrooms than in random grouped classrooms. Because both grouping situations provide considerable need-fulfillment for the high achieving pupil, it was expected that the largest differences would be found between slow pupils in the two districts. No differences between average pupils in the two grouping treatments were hypothesized because the curriculum in random grouped classrooms is usually well adapted for average pupils, thus suggesting that little real difference between the treatments exists at this ability level.

Since the ability grouped classroom would appear to adapt to the needs of pupils at all three levels with about equal effectiveness, no significant within-treatment differences in attitude towards the teacher were hypothesized for District A. In District R, the average pupil would seem to have the fewest frustrations because the rate and level of teaching is well adapted to his abilities. The high achieving pupil, however, also experiences considerable need-fulfillment in the random grouped classroom even though the curriculum may be geared at a rate below his capacity. He receives considerable recognition and also can gain satisfaction from working on enrichment materials available. Thus, it was hypothesized that superior and average pupils in District R would exhibit similar attitudes towards the teacher while the attitudes of slow pupils would be significantly less favorable.

There seems to be some evidence that girls tend to have generally more favorable attitudes towards the total school situation than boys. Thus, it was hypothesized that for both districts significant differences favoring girls would occur in the between-sex comparisons on this scale.

### Attitude Towards School

Most of the same variables discussed regarding attitudes towards the teacher seemed likely to operate in the same fashion on the pupil's attitudes towards school. Thus, it was hypothesized that no significant differences would occur in comparisons between average pupils in the two districts. It was further hypothesized that both superior and slow pupils in District A would have significantly more favorable attitudes towards school than comparable pupils in District R. No significant within-treatment differences were hypothesized for District A pupils in this attitude area. It was hypothesized that for District R pupils, average and superior pupils would both have more favorable attitudes towards school than slow pupils.

Regarding sex differences, it was hypothesized that girls' groups in both districts and at all levels would have more favorable attitudes towards school than comparable boys' groups, with the largest differences occurring at the superior level.

The My School scales on the SRA Youth and Junior Inventories are concerned with problems of school adjustment generally related to the attitude items contained on the USU School Inventory. Such items as "I dislike my courses," "I doubt the value of the things I study," "My teachers aren't interested in me," demonstrate the similarity between the problems contained in this scale and some of the inventory items. Our working hypothesis regarding this scale was that superior and slow District A pupils would report fewer problems than comparable pupils in District R. It was hypothesized that no differences in mean problems scores would be found between average pupils in the two districts.

The California Test of Personality, elementary level, was administered to Sample IV pupils during their fifth grade year. The School Relations scale of this test attempts to identify the pupil who is satisfactorily adjusted to his school, feels his teachers like him, and enjoys satisfactory peer relationships. The reader will note that this scale attempts to measure substantially the same areas as the USU School Inventory. This section of the California Test of Personality, however, contains only 12 items as compared with 95 items on the USU School Inventory. It seems likely that because of the somewhat more thorough coverage provided by the latter instrument, the results of this instrument might be accepted with somewhat more confidence. The working hypotheses established for this scale hypothesized more favorable scores for District A at the superior and slow level with no differences hypothesized for District R. No within-treatment differences were hypothesized for District A, while differences favoring superior pupils were hypothesized for District R.

## PREVIOUS RESEARCH RELATED TO PUPIL ATTITUDES AND GROUPING

Almost no research has been reported that is directly concerned with pupil attitudes toward school in ability grouped and random grouped classrooms. Some data in this area were collected by Simpson and Martinson (1961). The reader will recall that this study was concerned with the effectiveness of special programs for gifted children in California. Rating scales were developed for use by pupils, teachers, and parents that dealt with a number of areas somewhat related to pupil attitudes. Among the variables covered on these questionnaires were included interest in school and rapport with teachers. These data, however, were not collected for the control groups so that comparisons were not possible for most of the samples involved in the study. The fifth and sixth grade experimental programs, however, included both an enrichment program and a special class program. The first of these programs employed enrichment for superior pupils in the regular classroom, and the second involved a special class for superior pupils. These two programs were quite similar to the random grouping and ability grouping systems in Districts A and R of the Utah study. Rating scale data from pupils and parents were collected on these two groups at the beginning and end of the one year study, and percentage of gains was calculated.

The rating scale items on interest in school, motivation towards learning, and intellectual curiosity were classified together as "intellectual stimulus" by the investigators. Ratings by pupils and their parents in the special class situation indicated an average pupil gain on these items of about 80 per cent over the year of the study. Similar ratings for the enrichment program indicated gains of about 50 per cent.

On another pertinent variable, "relationships with the teacher," ratings by special class pupils and their parents showed a gain of about 45 per cent as compared with a gain of approximately 35 per cent for the enrichment program. Teacher ratings were not available for the special class groups. As each of the rating scales included only a single item for each area covered, these data are subject to serious limitations in reliability. The validity of a direct rating scale when used by presumably untrained raters must also be questioned.

Goldberg and her associates reported that in their study of the effect of ability grouping per se, comparisons between pupils in homogeneous and heterogeneous classes on attitudes towards school and attitudes towards more or less able peers yielded results that were largely negative.[1]

---

[1] Quoted from a personal communication from Dr. Goldberg, dated September 27, 1963, with the author's permission.

A study by Passow, Goldberg, and Link (1961) compared four groups of gifted junior high school pupils in mathematics achievement from the beginning of the seventh to the end of the ninth grade. Two of these groups participated in an accelerated program, one in an enrichment program and there was one control group. In addition to measuring achievement differences, this study administered an attitude inventory at the beginning of the seventh and end of the ninth grade. This inventory contained items in six attitude areas related to mathematics. The groups were closely comparable on the initial attitude test. Analysis of the total score of the attitude inventory administered at the end of the ninth grade showed the two classes in the accelerated program to have similar means (60.2 and 63.6). These means were significantly higher than those of pupils in the enrichment group (53.6) and the control group (53.7). These data suggest that acceleration develops considerably more favorable attitudes toward mathematics than enrichment in groups of homogeneously grouped gifted pupils. A limitation of this study is that only one classroom was involved in each of the four grouping treatments. Four different teachers taught these classes. Thus, it is very possible that some of the differences obtained were due to teacher differences rather than to treatment differences.

Drews' study (1962) of homogeneous and heterogeneous grouping in ninth grade English classes was also concerned with pupil attitudes. Interest in school was obtained by asking pupils to check "yes" or "no" in answer to the question, "Is school work interesting to you?" These data were obtained at the beginning and end of the one year study. The percentage of favorable responses of homogeneously grouped pupils at the beginning of the study was 80 per cent, and at the end of the study, 74 per cent. Comparable percentages for the heterogeneously grouped pupils were 78 per cent and 80 per cent. These data indicate that the two grouping treatments had no effect on the pupils' stated interest in school. The use of a single direct question to measure pupil interest raises serious doubt about the reliability and validity of these data.

Attitude towards English as a school subject was also obtained at the beginning and end of the research project by asking students to list school subjects they "liked most." The results showed that attitudes towards English improved for all groups excepting the heterogeneous slow group between the pre-testing and post-testing.

Four questions were asked of pupils in the two treatments to see if differences had developed between the attitudes of pupils in the homogeneous vs. the heterogeneous English classes. The results reflected about the same attitudes for average pupils in the two treatments, but for superior and slow pupils the homogeneous classes elicited much more favorable attitudes. For superior pupils, 92 per cent of those in homogeneous classes expressed favorable attitudes as opposed to only 33 per cent in heterogeneous classes. For slow pupils, 83 per cent of those in homogeneous groups expressed favorable attitudes as compared with 60 per cent of those in heterogeneous groups. Although tests for significance of these data were not presented, it is apparent that superior and slow pupils in homogeneous experimental classes in ninth grade English developed considerably more favorable attitudes than did comparable pupils in the heterogeneous classes.

## PROCEDURE

### Sample

Data relating to school attitudes were obtained for Samples IV, VI, VII, and VIII. The USU School Inventory, which was the most reliable tool employed in this area, was administered only to Sample IV. Sample IV pupils also were administered the California Test of Personality and the SRA Juhior Inventory and Youth Inventory. Data obtained from these measures that are not directly pertinent to school attitudes will be found discussed in Chapters 9 and 12.

All samples also were administered either the California Study Methods Survey or the Brown-Holtzman Survey of Study Habits and Attitudes at some time during the project. These measures, although related to this chapter, are treated in Chapter 6 which deals exclusively with these study methods and attitudes.

### Measures

#### The USU School Inventory

As a measure of school attitudes meeting the requirements of this study could not be found, it was decided to develop such a scale. A large number of statements expressing various ideas and feeling about school-related topics were collected from both teachers and pupils. These statements were reviewed and rewritten in a question form that permitted answering "yes," "no," or "?." Of the statements collected, 109 were selected for inclusion in the pretest. This pretest was administered to 130 sixth grade pupils from four elementary schools in a Utah school district similar to the Districts A and R. Data collected showed these pupils to be comparable in socioeconomic background to the Sample IV pupils for whom this measure was developed.

The pretest results were item-analyzed and item validities were computed using the Flanagan tables (1939). Item validity coefficients generally ranged from .45 to .70, although a few items on each scale were outside of this range. A coefficient of .30 was established as the minimum cutoff for item validity. Only one of the 25 items on the Attitudes Towards Peers scale, one item of the 20 relating to Attitude Towards School, and two items of the 50 included on the Attitude Towards the Teacher scale had item validities below .40.

The reliability coefficient for the 95 items on the total scale was .94, using the split-half correlation corrected for length with the Spearman-Brown Prophesy Formula.[1]

Product moment correlation coefficients were also calculated between scores on the three subscales to determine their level of independence. The correlation coefficient between school attitude and peer attitude was .54; between attitude towards teacher and towards peers was .46; and between attitude towards the teacher and towards school was .57. In view of the fairly high reliabilities of the subscales, these intercorrelations suggest that the three subscales, though related, are measuring somewhat different facets of the pupil's attitude.

In order to give the reader further insight into the attitudes measured by the USU School Inventory, Table 8-1 lists the five items having the highest item validity from each of the three subscales. Complete item analysis data for this instrument may be obtained from the investigators. A copy of this measure may be found in Appendix D.

---

[1] Reliabilities of the subscales were .95, .91, and .82.

Table 8-1

Sample Items From the USU School Inventory

| Items | Item Validity |
|---|---|
| **Attitude Toward Peers:** | |
| 37. Are there any groups of children who won't let you play with them? | .68 |
| 61. Do you think that the boys and girls in this school like you as well as they should? | .68 |
| 64. Do your classmates usually feel that they know more than you? | .68 |
| 68. Are you often left out of things other kids do? | .73 |
| 75. Do you find it hard to be as popular as the other kids? | .73 |
| **Attitude Toward the Teacher:** | |
| 28. Do you feel that some of your teachers have held a "grudge" against you? | .77 |
| 81. Is your teacher "bossy"? | .77 |
| 85. Do you often think that your teacher does not like you? | .81 |
| 91. Do you often wish you had some other teacher? | .79 |
| 93. Is your teacher often so unkind or unfair that it makes you feel bad? | .79 |
| **Attitude Toward School:** | |
| 1. Do you like all of the subjects you are now taking in this school? | .77 |
| 33. Do you have trouble keeping your mind on what you are studying? | .77 |
| 46. Are you smart in school? | .79 |
| 48. Are some of your lessons very boring to you? | .79 |
| 74. Do you like your lessons? | .73 |

In addition to the USU School Inventory, the My School scales on the SRA Youth and Junior Inventories and the School Relations score on the California Test of Personality were also administered.

## RESULTS

The results of the attitude study were analyzed using a single classification analysis of variance to compare attitude scores of pupils in the two grouping treatments, at different ability levels and of different sex. Pupils in each district were divided into six groups on the basis of ability level and sex. Comparisons between pupils of the same ability level classification and sex were made to appraise the differential effects of the two treatments. Comparisons between pupils of the same sex but different ability levels within each district were carried out to determine the effects of each treatment on pupils of different ability. Comparisons between pupils in the same treatment and ability level but of different sex were carried out to determine the nature and extent of sex differences on each variable.

### The USU School Inventory

#### Attitude Towards Peers

Analysis results of the between-treatment comparisons on the Attitude Towards Peers scale may be found in Table 8-2. It will be noted that none of these differences approached significance nor did the slight differences found occur consistently in favor of one district or the other. These results would indicate that the two grouping treatments had no significant effect upon the attitudes measured on this portion of the USU School Inventory. Within-treatment comparisons on the Attitude Towards Peers scale generally supported the working hypotheses. Four of the six differences between pupils of different ability levels in District R were statistically significant with the higher ability group obtaining the more favorable attitudes in all cases. Within-treatment comparisons for District A pupils revealed much smaller differences with only one of the six differences reaching statistical significance at the .05 level. These data suggest that attitudes of pupils towards their peers are rather consistently related to ability in random grouped classes. It is noteworthy in reviewing these differences in Table 8-3 that while average boys had significantly more favorable attitudes than developmental boys in District R, the differences in District A were slight with developmental boys actually having a slightly higher mean score.

Results of the analysis of sex differences in the Attitude Towards Peers score may be found in Table 8-4. Sex differences were slight on this variable. Girls were slightly higher in four of the five comparisons, but the only difference approaching statistical significance favored average boys over average girls in the District R sample. These results would suggest that no real sex differences were present on the scores in this attitude area.

#### Attitude Towards the Teacher

The pupil's attitude toward his teacher was considered the most important of the three areas that the USU School Inventory was designed to measure. Of the 95 items on the inventory, 50 were aimed at measuring this aspect of pupil attitude. It was

Table 8-2

Between-Treatment Differences on the USU School Inventory

| Group | $N_A$ | $N_R$ | Attitude Toward Peers<br>Mean A | Attitude Toward Peers<br>Mean R | Attitude Toward Peers<br>t | Attitude Towards the Teacher<br>Mean A | Attitude Towards the Teacher<br>Mean R | Attitude Towards the Teacher<br>t | Attitude Towards School<br>Mean A | Attitude Towards School<br>Mean R | Attitude Towards School<br>t |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Superior Boys<br>Dist. A vs. R | 48 | 138 | 16.13 | 16.61 | NS[a] | 43.14 | 40.06 | 2.04 | 12.38 | 12.32 | NS |
| Superior Girls<br>Dist. A vs. R | 47 | 143 | 16.79 | 16.94 | NS | 45.74 | 40.82 | 3.26 | 14.38 | 13.74 | NS |
| Average Boys<br>Dist. A vs. R | 83 | 98 | 14.66 | 15.90 | NS | 37.39 | 37.46 | NS | 180.80 | 11.56 | NS |
| Average Girls<br>Dist. A vs. R | 88 | 108 | 15.32 | 14.61 | NS | 37.87 | 38.30 | NS | 11.37 | 11.56 | NS |
| Slow Boys<br>Dist. A vs. R | 31 | 88 | 15.10 | 13.98 | NS | 41.64 | 36.63 | 2.66 | 12.23 | 10.09 | 2.68 |
| Slow Girls<br>Dist. A vs. R | 14 | 42 | 13.36 | 14.36 | NS | 39.21 | 38.83 | NS | 11.78 | 12.19 | NS |

[a] Not statistically significant.

## Table 8-3

Within-Treatment Differences on the USU School Inventory

| Group | Attitude Towards Peers |  |  | Attitude Towards the Teacher |  |  | Attitude Towards School |  |  |
|---|---|---|---|---|---|---|---|---|---|
|  | Mean 1 | Mean 2 | t | Mean 1 | Mean 2 | t | Mean 1 | Mean 2 | t |
| **District A** |  |  |  |  |  |  |  |  |  |
| (1) Superior Boys vs.<br>(2) Average Boys | 16.13 | 14.66 | NS[a] | 43.14 | 37.39 | 3.51 | 12.38 | 10.80 | 2.26 |
| (1) Superior Boys vs.<br>(2) Slow Boys | 16.13 | 15.10 | NS | 43.14 | 41.64 | NS | 12.38 | 12.23 | NS |
| (1) Average Boys vs.<br>(2) Slow Boys | 14.66 | 15.10 | NS | 37.39 | 41.64 | 2.24 | 10.80 | 12.23 | AS[b] |
| (1) Superior Girls vs.<br>(2) Average Girls | 16.79 | 15.32 | NS | 45.74 | 37.87 | 4.83 | 14.38 | 11.37 | 4.30 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 16.79 | 13.36 | 2.09 | 45.74 | 39.21 | 2.37 | 14.38 | 11.78 | 2.20 |
| (1) Average Girls vs.<br>(2) Slow Girls | 15.32 | 13.36 | NS | 37.87 | 39.21 | NS | 11.37 | 11.78 | NS |
| **District R** |  |  |  |  |  |  |  |  |  |
| (1) Superior Boys vs.<br>(2) Average Boys | 16.61 | 15.90 | NS | 40.06 | 37.46 | 2.18 | 12.32 | 11.60 | NS |
| (1) Superior Boys vs.<br>(2) Slow Boys | 16.61 | 13.98 | 3.60 | 40.06 | 36.63 | 2.79 | 12.32 | 10.09 | 4.21 |
| (1) Average Boys vs.<br>(2) Slow Boys | 15.90 | 13.98 | 2.43 | 37.46 | 36.63 | NS | 11.60 | 10.09 | 2.65 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (1) Superior Girls vs.<br>(2) Average Girls | 16.94 | 14.61 | 3.42 | 40.82 | 38.30 | 2.19 | 13.74 | 11.56 | 4.45 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 16.94 | 14.36 | 2.74 | 40.82 | 38.83 | NS | 13.74 | 12.10 | 2.99 |
| (1) Average Girls vs.<br>(2) Slow Girls | 14.61 | 14.36 | NS | 38.30 | 38.83 | NS | 11.56 | 12.10 | NS |

[a] Not statistically significant.
[b] Approached statistical significance.

hypothesized that the scores of superior and slow pupils in District A would reflect significantly more favorable attitudes than comparable District R pupils and that no differences would be present between average pupils in the two districts. It will be noted in Table 8-2 that both superior groups in District A scored significantly higher on this subscale than comparable District R pupils. Average pupils in the two districts received very similar mean scores with differences falling far short of statistical significance. Slow boys in District A received significantly more favorable scores than comparable District R pupils but slow girls were not significantly different in the two districts. Thus, the working hypotheses were well supported for superior and average pupils and partially supported for slow pupils.

Comparisons between pupils of different ability levels on the Attitude Towards the Teacher score revealed three statistically significant differences in the six comparisons made for District R pupils. These differences all involved superior groups and indicated a tendency for superior pupils in District R to have somewhat more favorable attitudes than either average of slow pupils. In District A no significant within-treatment differences were hypothesized. The results, however, indicate four significant differences on the six comparisons. Three of the four comparisons involving superior students were significant, in all cases favoring the superior group. The fourth significant difference was between average and slow boys and favored the slow group. Slow girls in District A showed more favorable attitudes towards the teacher than did average girls although this difference did not reach statistical significance. It appears that the significant difference favoring slow boys is particularly noteworthy as it reverses the pattern found in most of the District R data that showed favorable attitudes to be directly related to ability level.

It was hypothesized that significant sex differences favoring girls would be found in both districts on the Attitude Towards the Teacher score. As may be seen in Table 8-4, no significant sex differences emerged on this variable although five of the six non-significant differences did occur in the hypothesized direction.

### Attitude Towards School

It was hypothesized that Attitude Towards School scores would favor District A pupils in comparisons at the superior and slow ability levels and that no differences would occur between average pupils in the two districts. It may be seen in Table 8-2 that only one significant difference emerged. This difference was between slow boys in the two districts and favored District A pupils at the .01 level. It appears likely that the hypothesized differences between superior pupils did not emerge because of the fact that superior (high achieving) pupils in both districts obtained considerable fulfillment of their psychological needs from their school work. Superior District A pupils did receive slightly higher scores but these differences fell far short of statistical significance.

It was hypothesized that no within-treatment differences on Attitude Towards School scores would be found in District A. It may be seen in Table 8-3 that three significant differences occurred, all favoring superior pupils as compared with pupils of average or low ability. These differences in all likelihood are a reflection of the need-fulfillment provided in both districts for high achieving pupils. Within-treatment comparisons in District R showed a similar pattern with four significant differences emerging all of which favored the group of higher ability.

It was hypothesized that girls would have significantly more favorable attitudes towards school in both districts than boys of comparable ability. Three of the six

Table 8-4

Sex Differences on the USU School Inventory

| Group | Attitude Towards Peers | | | Attitude Towards the Teacher | | | Attitude Towards School | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean B | Mean G | t | Mean B | Mean G | t | Mean B | Mean G | t |
| District A | | | | | | | | | |
| Superior Boys vs. Superior Girls | 16.13 | 16.79 | NS[a] | 43.14 | 45.74 | NS | 12.38 | 14.38 | 2.53 |
| Average Boys vs. Average Girls | 14.66 | 15.32 | NS | 37.39 | 37.87 | NS | 10.80 | 11.37 | NS |
| Slow Boys vs. Slow Girls | 15.10 | 13.36 | NS | 41.64 | 39.21 | NS | 12.23 | 11.78 | NS |
| District R | | | | | | | | | |
| Superior Boys vs. Superior Girls | 16.61 | 16.94 | NS | 40.06 | 40.82 | NS | 12.32 | 13.74 | 3.09 |
| Average Boys vs. Average Girls | 15.90 | 14.61 | NS | 37.46 | 38.30 | NS | 11.60 | 11.56 | NS |
| Slow Boys vs. Slow Girls | 13.98 | 14.36 | NS | 36.63 | 38.63 | NS | 10.09 | 12.19 | 2.92 |

[a] Not statistically significant.

comparisons between the sexes yielded statistically significant differences, all in the hypothesized direction. These data provide fairly clear evidence that girls may be expected to show more favorable attitudes toward school than boys, especially at the superior level.

## School Problems

The My School scales of the SRA Junior Inventory and Youth Inventory are concerned with problems of school adjustment many of which appear to reflect attitudes similar to those measured on the USU School Inventory. Results of the analysis of this scale for the four samples to which one of the SRA inventories was administered in the Utah study may be found in Table 8-5. It will be noted that in the six comparisons between Sample IV pupils in the two districts, District R superior and average girls had significantly lower problem scores than comparable District A pupils. No significant differences were found in the Sample VI or Sample VII comparisons. Sample VIII comparisons, however, revealed two significant differences both favoring District A pupils and a third difference approaching significance also favored the District A sample. Thus, of a total of four significant differences found over the four samples, two favored District R pupils and two favored District A pupils. No consistent trend appeared to be present in the data on this scale suggesting that the two treatments had produced no differential effects upon the number of school problems pupils reported.

## School Relations

Results on the School Relations scale of the California Test of Personality may be found in Table 12-5 (page 350). It will be noted that of the six between-treatment comparisons, three were statistically significant at the .05 level. District R superior boys and girls both received higher scores on this instrument as did District R slow boys. These differences, which generally contradicted the findings of the USU School Inventory, were collected during the fifth grade. It is not possible to determine whether the disagreement between the two measures indicates that they are measuring different factors, or reflects a later improvement in school attitudes of District A pupils as shown on the USU School Inventory. In the case of conflicting results of this sort, it seems most logical to accept the findings of the more reliable instrument which in this case is the USU School Inventory. Within-treatment differences were hypothesized to favor pupils of higher ability in District R and to be nonsignificant in District A. Pupils of higher ability in both districts generally received more favorable scores on this instrument. Four of the six differences were statistically significant in District A and three of these differences were significant in District R (see Table 12-6).

It was hypothesized that girls would receive more favorable scores in both districts than boys on this instrument. This trend, though slight, was present with girls receiving slightly higher scores on all six of the between-sex comparisons. Two of these differences, between superior boys and girls in the two districts, were statistically significant.

## Table 8-5

### Between-Treatment Differences on the SRA My School Scale

| | Sample IV<br>SRA Jr. Inv.<br>5th Grade | | | Sample VI<br>SRA Youth Inv.<br>7th Grade | | | Sample VII<br>SRA Youth Inv.<br>8th Grade | | | Sample VIII<br>SRA Youth Inv.<br>9th Grade | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Mean A | Mean R | t | Mean A | Mean R | t | Mean A | Mean R | t | Mean A | Mean R | t |
| Superior Boys<br>Dist. A vs. R | 30.57 | 28.74 | NS | 7.65 | 7.79 | NS | 6.32 | 7.62 | NS | 6.32 | 8.09 | 2.32 |
| Superior Girls<br>Dist. A vs. R | 28.28 | 23.43 | 2.26 | 6.95 | 6.76 | NS | 8.83 | 6.17 | NS | 6.65 | 7.22 | NS |
| Average Boys<br>Dist. A vs. R | 36.91 | 33.85 | NS | 7.59 | 8.62 | NS | 8.59 | 9.62 | NS | 9.00 | 9.58 | 2.07 |
| Average Girls<br>Dist. A vs. R | 35.31 | 28.44 | 3.15 | 8.10 | 7.73 | NS | 9.08 | 10.63 | NS | 7.93 | 9.27 | NS |
| Slow Boys<br>Dist. A vs. R | 39.38 | 40.14 | NS | 10.31 | 9.15 | NS | 10.11 | 10.54 | NS | 8.77 | 11.55 | AS |
| Slow Girls<br>Dist. A vs. R | 27.75 | 36.13 | NS | 9.55 | 11.06 | NS | 12.29 | 10.20 | NS | 8.77 | 10.35 | NS |

## SUMMARY AND CONCLUSIONS

Pupils in Sample IV were administered the USU School Inventory during their sixth grade year. The USU School Inventory contains three scales: Attitude Towards Peers, Attitude Towards the Teacher, and Attitude Towards School. Sample IV pupils were also administered the California Test of Personality and the SRA Junior Inventory, both of which contain scales that appear related to pupils' attitudes toward school. Pupils in Sample VI, VII, and VIII were administered the SRA Youth Inventory. One of the problems areas covered on this measure also appears to be related to school attitudes. Comparisons between the two grouping treatments, between pupils of different ability levels within each grouping treatment, and between comparable pupils of the opposite sex were made using single classification analysis of variance. Sigfigicant differences between mean scores of various subgroups on the áforementioned measures form the basis for our conclusions.

Results on the Attitude Towards Peers scale of the USU School Inventory revealed no significant between-treatment differences. Four of the six within-treatment differences in District R were significant, favoring the higher ability group in each case. No significant sex differences were found on this scale. These results suggest the following conclusions:

1. The two grouping treatments had no significant effect on attitudes of pupils toward their peers.

2. Attitude towards peers was found to be consistently related to ability in the random grouped classrooms while no such relationship was found in the ability grouped classrooms, suggesting that ability is much more important in determining peer attitudes in the random grouped situation.

3. There was no relationship between sex and attitude towards peers.

Comparison of mean scores on the Attitude Towards the Teacher scale of the USU School Inventory, showed that superior girls and boys and slow boys in District A received significantly more favorable scores than comparable District R pupils. No significant differences were found between average pupils in the two districts. Within-treatment comparisons between pupils at different ability levels showed three significant differences in District R, all favoring the superior groups. There were four significant differences between District A pupils of different ability levels, three favoring superior groups and one favoring slow boys over average boys. No significant sex differences were found in this attitude area. From these data it may be concluded that:

4. Ability grouping appears to cause more favorable attitudes towards the teacher than random grouping for superior and slow pupils.

5. Superior pupils in both districts indicated more favorable attitudes towards the teacher than pupils at the lower ability levels.

6. There was some tendency in District A for slow pupils to have more favorable attitudes towards the teacher than average pupils, thus supporting the conclusion that ability grouping has a favorable effect upon the attitudes of slow pupils towards their teacher.

7. There were no differences between the mean attitudes of comparable boys' and girls' groups toward the teacher.

Comparisons between the two districts on the Attitude Towards School scale showed one significant difference favoring slow boys in District A. In both districts there were significant differences between pupils of different ability levels, all indicative of more favorable attitudes for pupils of higher ability. Three of the six sex differences on the Attitude Towards School scale were statistically significant. These differences, all of which favored girls, were largest at the superior level. It may be concluded that:

8. Boys of low ability developed more favorable attitudes toward school in ability grouped classroom than in random grouped classrooms.
9. Superior pupils had more favorable attitudes toward school in both treatments than pupils of lower ability.
10. Superior girls had more favorable attitudes towards school in both districts than boys of comparable ability.

## REFERENCES

Drews, E. M. The effectiveness of homogeneous and heterogeneous ability grouping in ninth grade English classes with slow, average, and superior students. Unpublished manuscript. Michigan State University, 1962.

Flanagan, J. C. General considerations in the selection of test items and a short method of estimating the product-moment coefficient from the data at the tails of a distribution. J. educ. Psychol., 1939, 30, 674-80.

Passow, A. H., Goldberg, M. L., & Link, F. R. Enriched mathematics for gifted junior high school students. Educ. Leadership, 1961, 18, 442-448.

Simpson, R. E., & Martinson, R. A. Educational programs for gifted pupils. Sacramento, California: California State Department of Education, January, 1961.

## Chapter 9

# PUPIL PROBLEMS

## INTRODUCTION

The main purpose of this chapter is to compare the problems reported by pupils in ability grouped and random grouped classes. The instruments employed were the SRA Youth Inventory and the SRA Junior Inventory. These measures are essentially check lists each of which classifies the problems listed into a number of areas. A study of these problem areas on the two inventories revealed several areas in which one of the grouping treatments employed in the Utah study might be expected to lead to greater numbers of pupil problems than the other grouping treatment. Two of the areas covered, those concerned with school problems and self-perception, appeared likely to yield information that would further test some of the hypotheses that were already being tested by other instruments in the Utah study and would thus provide a check of these results. Such checking of results is always desirable and is especially valuable in areas such as self-perception because of the relative newness of the related psychological constructs and the absence of established measuring instruments.

## WORKING HYPOTHESES

As no previous research has been reported in this area, working hypotheses were established by weighing the possible psychological effects of ability grouping and random grouping on the various problem areas covered on the SRA Junior Inventory. No hypotheses were established with regard to within-treatment differences, but these data were analysed and reported in order to provide some insight into the problems encountered by pupils of different ability levels in the two treatments.

The "Looking Ahead" area of the SRA Youth Inventory is concerned primarily with problems related to choosing a career, going to college, and finding a job (Remmers, et. al, 1953). Many of the problems included in this section are concerned with the pupil's insight into his ability. It was thus hypothesized that District A pupils would indicate fewer problems in this area than District R pupils at all ability levels with greatest differences occurring between superior and average subgroups in the two districts.

The next problem area to be discussed, "Getting Along with Others," is found in both of the SRA inventories. In the Junior Inventory, the majority of the items relate to social adjustment problems, while in the Youth Inventory, social skills problems are emphasized (Remmers and Bauernfeind, 1955). It was hypothesized, however, that District R pupils would have a greater number of problems on both scales than the comparable District A pupils. Because of the reduced chances of social rejection stemming from ability differences, it would seem that social adjustment problems would be less serious in District A, especially for pupils of low ability. Social skills appear less likely to be influenced by the grouping treatments, but again it appeared probable that the greater opportunities in District A for pupils at the average and slow levels to gain leadership positions and play more active roles in social activities, would tend to result in a lower frequency of social problems.

The "Boy Meets Girl" area of the SRA Youth Inventory is concerned with dating and sex information, and sex mores. It seemed unlikely that the grouping treatments would have any effect on the latter two topics. With regard to dating practices, however, it appeared likely that District A pupils would report fewer problems. Ability grouping at all levels would appear to place pupils in contact with more individuals having similar interests and aspirations. Also, the more favorable social climate present in the average and slow ability grouped sections should tend to reduce dating problems. Thus, it was hypothesized that between-treatment differences in the "Boy Meets Girl" problem area would generally favor pupils in ability grouped classes.

"The Basic Difficulty Key" for the SRA Youth Inventory is made up of those items considered by mental hygiene experts to be indicative of possible personality disturbance (Remmers, et. al., 1953). Although our working hypotheses on the SRA problem areas generally hypothesized more problems for pupils in random grouped classes, large differences were not expected, and there seemed to be little reason to expect either grouping system to bring about serious personality disturbances. Thus, it was hypothesized that no significant between-treatment differences would occur on the "Basic Difficulty Key."

Three of the problem areas found in the SRA Junior and Youth Inventories were administered only because removing them from the instrument would have exposed subjects to a non-standard administration of the measure that could in turn have changed responses in some of the areas where directional hypotheses were established. No working hypotheses were developed for these areas. These three problem areas were: "Home and Family," "Health," "Things in General." Although no hypotheses were established for these areas, it was thought that the results might be of interest to some readers and for that reason these data were analyzed and results are reported in this chapter.

No working hypotheses were established for sex differences on the SRA Junior Inventory or Youth Inventory, but data were analyzed for sex differences and findings are reported.

## PROCEDURE

### Subjects

The SRA Junior Inventory, form S was administered to Sample IV pupils in District A and R during their fifth grade year. The SRA Youth Inventory, form A was administered to Sample VI pupils during their seventh grade year, Sample VII pupils during their eighth grade year, and Sample VIII pupils during their ninth grade year. It will be noted that for all samples, these data were collected during the second year of the study, after District A pupils had been exposed to about one and one-half years in the ability grouping situation.

## Measures

The problem areas in the SRA Junior Inventory include the following: "About Me and My Home," "About Myself," "Getting Along With Other People," and "Things in General." The SRA Junior Inventory, form S provides a means for the pupil to indicate the seriousness of each problem. In addition to counting the number of problems checked in each of the aforementioned areas, a count of each pupil's "big problems" was made.

The SRA Youth Inventory, form A was administered to junior high school samples in grades seven, eight, and nine during the second year of the study. The Youth Inventory is similar to the Junior Inventory excepting for differences necessary in the Youth Inventory to better fit the problems of junior and senior high school pupils. On the Junior Inventory, form S the pupil may indicate the seriousness of his problems, while this cannot be indicated on the Youth Inventory, form A.

The Youth Inventory contains items in eight problem areas. These areas are entitled: "My School," "Looking Ahead," "About Myself," "Getting Along With Others," "My Home and Family," "Boy Meets Girl," "Health," and "Things in General." In addition to these eight problem areas, which contain a total of 298 items, a "Basic Difficulty Key" has been developed which provides another score on the Youth Inventory. This key was developed by classifying items in the Inventory into "Those which mental hygiene experts suggest may be indicative of a basic personality disturbance" and "Those which are essentially a matter of problem recognition" (Remmers, et. al., 1953, p. 4). The seven mental hygiene experts employed in making this classification, agree unanimously on 47 items indicative of a basic personality disturbance and six of the seven experts agreed on an additional 54 items. The "Basic Difficulty Key" is made up of these 101 items. A more detailed description of each problem area in the Youth Inventory and Junior Inventory is given along with the results on the areas, in the hope that placing these descriptive data with the results would help the reader make a more meaningful interpretation.

## RESULTS

### Looking Ahead

This area is not included in the Junior Inventory, as questions are concerned primarily with problems related to choosing a career, going to college and finding a job. Typical items in this section include: "How much ability do I actually have?"; "Shall I go to college?"; "Do I have the ability to do college work?"; "What jobs are open to high school graduates?" (Remmers and Bauernfeind, 1955).

Of the 18 between-treatment comparisons of group means in this problem area, there were nine differences that reached or approached statistical significance, all favoring the District A subgroups (see Table 9-1). As a number of the questions in this area are concerned with ability it is not surprising to note that the pupils in ability grouped classrooms reported fewer problems in this area than pupils in random grouped classrooms. This difference was found most commonly among superior boys and girls and average boys and may reflect a tendency for the District R pupils to have somewhat more doubts as to the level of their ability, thus leading to a larger number of problems.

Table 9-2 summarizes within-treatment differences on the mean number of problems in the "Looking Ahead" area. It will be seen that nine of these within-treatment

Table 9-1

Between-Treatment Differences on SRA Junior and Youth Inventories

| | | "Looking Ahead" | | | "Getting Along With Others" | | | | "Home and Family" | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Sample | VI | VII | VIII | IV | VI | VII | VIII | IV | VI | VII | VIII |
| Superior Boys | | R[a]<br>AS | R<br>AS | NS | NS | R<br>AS | R<br>1.96 | NS | NS | NS | NS | R<br>2.12 |
| Superior Girls | | R[b]<br>2.08 | NS | R<br>2.68 | NS | R<br>AS | NS | NS | NS | R<br>AS | NS | NS |
| Average Boys | | R<br>2.84 | R<br>2.07 | R<br>AS | A<br>2.29 | R<br>AS | NS | NS | A<br>2.02 | NS | NS | NS |
| Average Girls | | NS | NS | NS | A<br>3.18 | R<br>AS | NS | NS | A<br>2.74 | NS | R<br>AS | NS |
| Slow Boys | | NS | NS | R<br>2.67 | NS | NS | NS | R<br>2.55 | NS | NS | NS | NS |
| Slow Girls | | NS | NS | R<br>3.34 | NS | NS | NS | R<br>2.09 | NS | NS | NS | NS |

| | "Boy Meets Girl" | | | "Health" | | | "Things in General" | | | | "Basic Difficulty Key" | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample | VI | VII | VIII | VI | VII | VIII | IV | VI | VII | VIII | IV[c] | VI | VII | VIII |
| Superior Boys | R<br>2.34 | NS | R<br>3.22 | R<br>AS | NS | R<br>AS | NS | R<br>AS | NS | NS | NS | R<br>AS | NS | R<br>AS |
| Superior Girls | R<br>2.79 | NS | R<br>2.07 | NS | A<br>AS | NS | NS | R<br>2.15 | NS | NS | NS | NS | A<br>AS | R<br>AS |
| Average Boys | R<br>2.79 | NS | NS | R<br>2.47 | NS | NS | A<br>3.09 | R<br>3.22 | NS | NS | A<br>3.25 | NS | R<br>AS | NS |
| Average Girls | R<br>2.02 | R<br>2.17 | R<br>AS | NS | R<br>AS | NS | A<br>3.79 | NS | R<br>AS | R<br>2.66 | A<br>2.66 | NS | NS | R<br>2.70 |
| Slow Boys | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | R<br>AS |
| Slow Girls | NS | NS | R<br>2.03 | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS |

[a] Difference between indicated groups approaching significance favoring district indicated (R or A).
[b] Difference between indicated groups significant favoring district indicated and giving "t" value.
[c] Serious problems on Junior Inventory.

differences between pupils at different ability levels reached or approached statistical significance, six in District A and three in District R. These data reflect a tendency for pupils of higher ability to have lower mean problem scores.

Comparisons between scores of boys and girls in the "Looking Ahead" problem area revealed that girls consistently listed more problems in this area than boys. It may be seen in Table 9-3, however, that only one of the 18 sex differences reached the .05 level of significance, indicating that this trend, while consistent, is slight.

## Getting Along With Others

This section of the *Youth Inventory* is concerned with problems related to the pupil's social development with considerable emphasis upon social skills. Typical of the items in this area are: "I wish I could carry on a pleasant conversation"; "I want to learn to dance"; and "I need to learn how to plan a party." Social skill items make up about two-thirds of the 40 items in this area. Most of the remaining items in this section deal more directly with social adjustment, for example: "I wish I were more popular," and "I am not attractive to the other sex" (Remmers, *et*. *al*., 1953).

The *SRA Junior Inventory* contains 34 items in this area. In contrast to the *Youth Inventory*, the great majority of *Junior Inventory* items relate to social adjustment problems both with adults and peers, with very few pertaining directly to social skills. In view of these differences, Sample IV data are not directly comparable to data from Samples VI, VII, and VIII.

Between-treatment comparisons for Sample IV showed District A average pupils, both boys and girls, to have significantly higher mean problem scores in this area than comparable District R pupils. Although not significant, this same trend was found in the mean scores of slow boys and girls.

Results for Samples VI, VII, and VIII who were administered the *Youth Inventory*, showed a trend for District R pupils to have more problems in this area at all three ability levels. Of the 18 comparisons made at these grade levels three were statistically significant and four approached significance, all in the aforementioned direction. It will be noted in Table 9-1 that this trend is no more pronounced at one ability level than another.

The above data would indicate that at the elementary level, District A pupils listed more social adjustment problems than their District R counterparts, while at the secondary level, District R pupils listed more problems on the *Youth Inventory*, which it will be recalled deals more with social skills than social adjustment. The latter trend is consistant enough to indicate a true difference that is general at all three ability levels.

Within-treatment differences for Sample IV showed a consistant tendency for pupils of higher ability to list fewer social adjustment problems. These data are fairly consistant for both districts with four of the six District A comparisons and three of the six District R comparisons reaching statistical significance.

In contrast, the within-treatment comparisons at the secondary level show no consistent pattern. Of the 36 comparisons between groups at different ability levels in Samples VI, VII, and VIII, only two reached statistical significance at the .05 level, thus suggesting that no real ability level differences exist in frequency of problems involved with social skills.

Table 9-2

Differences Between Pupils at Different Ability Levels on SRA Junior and Youth Inventories

| Group | Sample | "Looking Ahead" | | | "Getting Along With Others" | | | | "Home and Family" | | | | "Boy Meets Girl" | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | VI | VII | VIII | IV | VI | VII | VIII | IV | VI | VII | VIII | VI | VII | VIII |
| **District A** | | | | | | | | | | | | | | | |
| (1) Superior Boys vs. | | NS | NS | NS | 2.29 | NS | NS | NS | NS | NS | NS | 2.14 | NS | NS | AS |
| (2) Average Boys | | | | | 2 | | | | | | | 2 | | | 2 |
| (1) Superior Boys vs. | | AS[a] | NS | NS | 3.20 | NS | NS | NS | 2.10 | NS | AS | NS | NS | NS | NS |
| (2) Slow Boys | | 2 | | | 2 | | | | 2 | | 2 | | | | |
| (1) Average Boys vs. | | AS | NS | NS | NS | NS | NS | 2.30 | NS | NS | NS | NS | NS | NS | NS |
| (2) Slow Boys | | 2 | | | | | | 1 | | | | | | | |
| (1) Superior Girls vs. | | 2.20[b] | NS | 2.74 | 3.23 | NS | NS | NS | 2.30 | 2.20 | NS | AS | AS | NS | NS |
| (2) Average Girls | | 2 | | 2 | 2 | | | | 2 | 2 | | 2 | 2 | | |
| (1) Superior Girls vs. | | 2.20 | NS | NS | 3.43 | NS | NS | NS | NS | AS | 2.54 | AS | NS | NS | NS |
| (2) Slow Girls | | 2 | | | 2 | | | | | 2 | 2 | 2 | | | |
| (1) Average Girls vs. | | NS | NS | 2.46 | NS | NS | NS | NS | NS | NS | 2.89 | NS | NS | NS | NS |
| (2) Slow Girls | | | | 1 | | | | | | | 2 | | | | |
| **District R** | | | | | | | | | | | | | | | |
| (1) Superior Boys vs. | | NS | NS | NS | NS | NS | NS | NS | NS | NS | AS | NS | NS | NS | NS |
| (2) Average Boys | | | | | | | | | | | 2 | | | | |
| (1) Superior Boys vs. | | NS | NS | 2.06 | 2.95 | NS | NS | NS | 2.71 | NS | 2.01 | NS | NS | NS | NS |
| (2) Slow Boys | | | | 2 | 2 | | | | 2 | | 2 | | | | |
| (1) Average Boys vs. | | NS | NS | 2.26 | NS | NS | NS | NS | AS | NS | NS | NS | NS | NS | NS |
| (2) Slow Boys | | | | 2 | | | | | 2 | | | | | | |
| (1) Superior Girls vs. | | NS | 2.00 | NS | NS | NS | 2.22 | NS | NS | NS | 2.24 | AS | NS | 3.47 | NS |
| (2) Average Girls | | | 2 | | | | 2 | | | | 2 | 2 | | 2 | |
| (1) Superior Girls vs. | | NS | AS | NS | 1.98 | NS | NS | NS | NS | NS | 2.90 | NS | NS | 2.14 | NS |
| (2) Slow Girls | | | 2 | | 2 | | | | | | 2 | | | 2 | |
| (1) Average Girls vs. | | NS | NS | NS | NS | NS | NS | NS | AS | NS | NS | NS | NS | NS | NS |
| (2) Slow Girls | | | | | | | | | 2 | | | | | | |

| Group | Sample | "Health" | | | "Things in General" | | | | "Basic Difficulty Key" | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | VI | VII | VIII | IV | VI | VII | VIII | IV[c] | VI | VII | VIII |
| District A | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | | NS | NS | NS | 2.35<br>2 | NS | NS | NS | 3.36<br>2 | NS | NS | 2.50<br>2 |
| (1) Superior Boys vs.<br>(2) Slow Boys | | NS | NS | NS | NS | 2.16<br>2 | NS | NS | 3.27<br>2 | AS<br>2 | AS<br>2 | NS |
| (1) Average Boys vs.<br>(2) Slow Boys | | NS | 2.40<br>1 | NS | NS | NS | NS | NS | NS | NS | AS<br>2 | NS |
| (1) Superior Girls vs.<br>(2) Average Girls | | NS | NS | NS | 3.54<br>2 | NS | NS | NS | 2.78<br>2 | AS<br>2 | NS | NS |
| (1) Superior Girls vs.<br>(2) Slow Girls | | NS | NS | NS | 2.46<br>2 | NS | NS | NS | 3.11<br>2 | NS | AS<br>2 | NS |
| (1) Average Girls vs.<br>(2) Slow Girls | | NS | NS | NS | NS | NS | NS | NS | NS | NS | AS<br>2 | NS |
| District R | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | | NS | NS | AS<br>1 | NS | NS | NS | NS | AS<br>2 | NS | NS | NS |
| (1) Superior Boys vs.<br>(2) Slow Boys | | NS | NS | NS | NS | NS | NS | NS | 4.21<br>2 | NS | 2.18<br>2 | NS |
| (1) Average Boys vs.<br>(2) Slow Boys | | NS | NS | NS | NS | NS | NS | NS | 2.47<br>2 | AS<br>2 | NS | NS |
| (1) Superior Girls vs.<br>(2) Average Girls | | NS | 3.51<br>2 | NS | NS | NS | 1.97<br>2 | NS | AS<br>2 | NS | 3.16<br>2 | 1.99<br>2 |
| (1) Superior Girls vs.<br>(2) Slow Girls | | NS | 2.29<br>2 | 1.96<br>2 | NS | NS | NS | NS | 4.72<br>2 | NS | 3.10<br>2 | 2.08<br>2 |
| (1) Average Girls vs.<br>(2) Slow Girls | | NS | NS | NS | NS | NS | NS | NS | 3.42<br>2 | NS | NS | NS |

---

[a] Indicates that difference between Sample IV Superior and Slow boys approached significance (AS) with group 2 (slow boys) having the greater number of problems.

[b] Indicates that difference between Sample IV Superior and Average girls was significant (t=2.20) with group 2 having the greater number of problems.

[c] Total number of serious problems compared for Sample IV, as SRA Junior Inventory does not have a Basic Difficulty Key.

Table 9-3

Sex Differences in Mean Problem Scores on the SRA Junior and Youth Inventories

| | | "Looking Ahead" | | | "Getting Along With Others" | | | | "Home and Family" | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Groups | Sample: | VI | VII | VIII | IV | VI | VII | VIII | IV | VI | VII | VIII |
| **District A** | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | | NS | NS | NS | NS | AS<br>G | 2.84<br>G | NS | NS | NS | NS | NS |
| Average Boys vs. Average Girls | | 2.03[a]<br>G | AS[b]<br>G | NS | NS | 3.23<br>G | 4.79<br>G | NS | NS | NS | NS | NS |
| Slow Boys vs. Slow Girls | | NS | NS | NS | NS | NS | 1.96<br>G | NS | NS | NS | AS<br>G | NS |
| **District R** | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS |
| Average Boys vs. Average Girls | | NS | NS | NS | NS | 2.06<br>G | 2.11<br>G | NS | NS | NS | NS | NS |
| Slow Boys vs Slow Girls | | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS |

|  | "Boy Meets Girl" |  |  | "Health" |  |  | "Things in General" |  |  |  | "Basic Difficulty Key" |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample: | VI | VII | VIII | VI | VII | VIII | IV | VI | VII | VIII | IV[c] | VI | VII | VIII |
| **District A** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | NS | NS | NS | NS | NS | AS<br>G | NS | NS | NS | NS | NS | NS | 2.78<br>G | NS |
| Average Boys vs. Average Girls | NS | 3.70<br>G | NS | 3.22<br>G | 3.70<br>G | AS<br>G | NS | NS | 2.25<br>G | NS | NS | 2.62<br>G | 4.81<br>G | NS |
| Slow Boys vs. Slow Girls | NS | NS | NS | NS | NS | NS | 2.05<br>G | NS | NS | NS | NS | NS | 2.12<br>G | NS |
| **District R** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS |
| Average Boys vs. Average Girls | NS | 3.04<br>G | NS | NS | 3.04<br>G | 3.03<br>G | NS | NS | AS | 2.68<br>G | NS | NS | 2.10<br>G | AS<br>G |
| Slow Boys vs. Slow Girls | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS | AS<br>G |

[a] Significant favoring group indicated, G for girls, B for boys, "t" value as given.

[b] Approached significance group indicated, G for girls, B for boys.

[c] Serious problems on Junior Inventory.

Sex comparisons in this area showed a consistent tendency in Samples VI and VII for girls to indicate more social skills problems than boys. In District R, two differences--between average boys and girls--reached statistical significance, and most of the non-significant differences also followed this pattern. The trend was somewhat more pronounced in District A where four of the sex difference comparisons reached statistical significance, and a fifth approached significance.

It appears likely that these differences are primarily a reflection of the greater interest in social skills displayed by girls during early and middle adolescence. Examination of the measure reveals items concerned with social skills that are considerably more important to girls than boys during the secondary school years. The higher scores made by girls may merely reflect the presence of such items.

## Boy Meets Girl

This area is not included in the SRA Junior Inventory. Most of the 32 items included in this area of the Youth Inventory are concerned with either dating, or sex information and mores with about equal coverage of these two topics. Typical items on dating include: "I don't know how to ask for a date," and "I don't know what to do on a date." Typical items on sex include: "I am embarrassed by any discussion on sex," and "I need more information about sex" (Remmers, et.al., 1953).

Review of between-treatment differences showed a consistent and marked trend for District R pupils to report more problems in this area than comparable pupils in District A (see Table 9-1). Eight of the 18 between-treatment comparisons were statistically significant and one additional approached significance; all in the aforementioned direction. Differences were generally larger and occurred most often between superior groups and least often between slow groups. Four of the six comparisons between superior pupils were statistically significant as compared with only one significant difference between slow pupils. Although between-treatment differences were hypothesized on this variable, the number of significant differences that emerged was somewhat surprising.

No differences between the districts were expected on items in this section related to sex information and mores. When a large number of significant between-treatment differences emerged, it was speculated that these differences might be due to a more extensive program of sex education in District A rather than being due to differences in dating problems, as was hypothesized. Therefore, to check this possibility, the sex education programs for the two districts were reviewed in order to determine whether any marked curriculum differences were present. Review of the activities of the junior high schools concerned indicated that relatively little time was devoted to sex education in either district prior to the ninth grade. Some specific problems such as menstration were dealt with briefly in the seventh and eighth grades, but the two districts appeared closely comparable in their treatment of these subjects. At the ninth grade level, junior high schools in District R devoted more time to sex education and related topics than did District A junior high schools. Thus, there seemed to be little in the curricula of the two districts to explain the rather consistent trend for District R pupils to report more problems in this area. If anything, the District R program gives the pupil more help in dealing with problems related to sex information than does the District A curriculum. Further research will be needed to determine whether the differences found in the Utah study are generalizable to other samples and if so to explore possible reasons why the ability grouped situation tends to be associated with fewer problems being reported on this section of the Youth Inventory.

Comparisons between pupils of different ability levels in each district revealed only two significant differences for District R and two differences approaching significance for District A (see Table 9-2). As a total of 36 within-treatment comparisons were made, it is questionable whether the four differences that reached or approached significance reflect any true difference. It should be noted, however, that all four of these differences indicate a greater number of problems reported by pupils of lower ability. Many of the nonsignificant differences reflected the same trend.

Of the 18 comparisons made between boys and girls on this area (Table 9-3) only two reached statistical significance, with average girls in both districts reporting more problems in this area than average boys.

The fairly marked trend in Table 9-1 in the direction of more problems for District R pupils appears likely to reflect a true between-treatment difference when we consider the few differences occurring within each district, the few sex differences occurring, and the similarity between the programs of the two districts in sex education.

## Home and Family

The "Home and Family" sections of the SRA Junior and Youth Inventories are concerned with problems related to parents, siblings, and the home environment. No differences between the two grouping treatments were hypothesized in this area. A review of between-treatment results in Table 9-1 shows District A average pupils at the fifth grade level to have reported significantly more problems on the Junior Inventory than comparable pupils in District R. For the three junior high school samples, however, only one significant difference emerged from the 18 comparisons on the Youth Inventory. It appears likely that this single significant difference was due to chance.

Within-treatment data on this problem area suggest a slight tendency for pupils of higher ability to report fewer problems. The eight between-sex comparisons in this area revealed no significant differences.

## Health

The SRA Youth Inventory contains 25 items dealing with health problems. No hypotheses were established for this scale. Of the 18 between-treatment comparisons, only one difference reached significance at the .05 level. As one significant difference in 20 might be expected by chance, it appears likely that this is a chance difference. Four other differences approached significance, three indicating a larger number of health problems in District R pupils than for District A pupils.

Of the 36 within-treatment comparisons made between mean health problem scores, only two reached statistical significance, again suggesting chance differences only.

A fairly clear trend of sex differences in this area appears to be reflected in Table 9-3. Of the 18 comparisons, four were significant beyond the .01 level and two others approached significance. All differences indicated a greater number of reported health problems for girls than boys of comparable ability. It is interesting to note that all four of the statistically significant differences were between average boys and girls, two in District A and two in District R. It is probably that the larger number of health problems checked by girls reflects the somewhat more dramatic physical changes experienced by girls during puberty. Some items on this scale were concerned with health problems relating to appearance, such as gaining or losing weight, improving the figure, body odor, and bad breath. In view of the greater tendency for girls to be preoccupied with social acceptance during the adolescent years, such items might be expected to cause more concern among girls than boys.

### Things in General

The "Things in General" sections of the *SRA Junior Inventory* and *Youth Inventory* differ considerably in emphasis. The problems included in the *Junior Inventory* deal primarily with children's interests, hobbies, and the development of manual and social skills. In contrast, this section of the *Youth Inventory* is concerned primarily with moral and religious problems. The reader should keep these differences in mind when comparing the results. It will be recalled that Sample IV pupils were administered the *Junior Inventory* and Samples VI, VII, and VIII, were administered the *Youth Inventory*.

Two significant between-treatment differences were found for Sample IV, indicating significantly more problems in this area for average boys and girls in District A. For Sample VI, superior boys and girls, and average boys in District R reported significantly more problems in this area. No significant differences were found, for Sample VII, and only one for Sample VIII. All of the four significant differences found on the *Youth Inventory* indicated higher (less favorable) scores for the District A pupils in the "Things in General" area but it is difficult to interpret this trend in view of the wide range of problems covered in this section. An item analysis, followed by a check of the results using another sample would probably be necessary to clarify these data.

### Serious Problems and Basic Difficulties

The *SRA Junior Inventory*, *form S* administered to Sample IV provides a means for pupils to identify the level of seriousness of their problems. The number of serious problems was obtained separately for pupils taking the *Junior Inventory* and these results are given in the last Sample IV columns in Tables 9-1, 9-2, and 9-3. Two significant between-treatment differences were found for Sample IV on this serious problem score, District A average boys and girls listed a greater average number of serious problems than comparable pupils of District R. Within-treatment differences on the number of the serious problems listed showed a marked and consistent trend in both districts for pupils of higher ability to indicate fewer serious problems. Four of the six within-treatment comparisons made in each district revealed statistically significant differences (see Table 9-2). No significant sex differences were found between comparable boys and girls in Sample IV on the serious problem score.

The "Basic Difficulty Key" of the *SRA Youth Inventory* includes 101 items taken from the *Inventory* that a panel of seven mental hygiene experts considered to be indicative of a possible basic personality disturbance. For the most part, the remaining items on the *Youth Inventory*, are more a matter of problem recognition. A number of comparisons of "Basic Difficulty" scores with external criteria reported in the test manual would indicate that the scale has some validity as a measure of overall adjustment (Remmers, *et*. *al*., 1953).

Between-treatment results showed that, although only one difference reached statistical significance, six other differences approached significance. All but one of these differences were found among superior and average students, and all but one difference indicated a higher basic difficulty score for pupils in District R. It is difficult to be sure on the basis of field study data that these differences came about as a result of the two grouping treatments. It does seem safe to conclude, however, that these findings refute the claim often made by educators that ability grouping tends to cause pupil maladjustment.

Within-treatment comparisons on the Basic Difficulty Scale showed a consistent tendency for pupils of higher ability to make lower (more favorable) scores. Of the 36 within-treatment comparisons, six were found to be statistically significant and seven others approached significance, all in the aforementioned direction. This trend generally held for both districts, although differences were somewhat larger in District R where five of the six statistically significant differences were found.

Of the 18 comparisons between boys' and girls' mean scores on the Basic Difficulty Scale, five reached statistical significance and two more approached significance. These differences reflect a fairly marked trend for girls to obtain higher Basic Difficulty scores than boys. This trend was more pronounced in District A where four of the nine sex differences were statistically significant, three beyond the .01 level (see Table 9-3).

## SUMMARY AND CONCLUSIONS

During the second year of the Utah study, Sample IV pupils were administered the SRA Junior Inventory, form S, and Sample VI, VII, and VIII pupils were administered the SRA Youth Inventory, form A. The results on these measures were analyzed using analysis of variance in order to compare the mean number of problems checked by pupils of comparable ability in the two districts, pupils at different ability levels in each district, and pupils of comparable ability but different sex in each district. As the SRA Youth Inventory was administered to three different samples at the junior high school level, an overall summary of the results on this measure seems desirable.

In reviewing the between-treatment differences on the problem areas of the Youth Inventory reported in this chapter, perhaps the most noteworthy trend is the preponderance of significant differences indicating larger mean problem scores for the District R groups. A total of 45 between-treatment differences reached or approached statistical significance for the junior high school samples. These differences were found in all samples: 18 in Sample VI, 10 in Sample VII, and 17 in Sample VIII. Of these 45 differences only two indicated higher mean problem scores for District A pupils while the other 43 indicated higher mean problem scores for District A. When the reader considers that these data are based upon pupils in three different junior high schools in each district, and when it is recalled that the samples were comparable on such variables as intelligence and socioeconomic status, the results would seem to suggest strongly that treatment differences between ability grouping and random grouping in mathematics and science may have a widespread effect upon the number of problems perceived by junior high school students. Whether this difference reflects a greater awareness of problems among District R pupils or reflects a tendency for ability grouping to reduce problems or help pupils deal with them cannot be answered from the data collected in this research. It should be noted, however, that the greatest between-treatment differences occurred in the problem areas in which it was hypothesized that differences would occur. Of the 45 differences reaching or approaching statistical significance, 23 were found in the "Looking Ahead," "Getting Along with Others" and "Basic Difficulty" scales. Of the four scales for which no differences were hypothesized, only the "Boy Meets Girl" scale showed marked and consistent between-treatment differences.

When the data in Table 9-1 are summarized by ability level another trend emerges--the tendency for the frequency of significant differences between the districts to be related to ability level. Of the 45 differences reaching or approaching statistical significance on the SRA Youth Inventory, 22 of these differences are found between superior pupils in the two districts, and in 20 of the 22 differences, the District R pupils obtained higher (less favorable) scores. Seventeen of these 45 differences were between average pupils in the two districts, with all differences indicating higher problem scores for District R groups. These results, reflecting as they do a consistently high number of problems for superior and average District R groups, would suggest the presence of true treatment differences at these ability levels. This trend, though present for slow pupils, is far less pronounced, and less general from sample to sample. Six differences were found between slow pupils in the two districts and all reflected higher problem scores for District R pupils. All six of the between-treatment differences found among slow pupils, however, occurred in Sample VIII, suggesting that they reflect unique characteristics of the sample rather than a generalizable treatment effect. These data support the overall conclusion that ability grouped junior high school pupils in average or superior groups report fewer problems than do comparable pupils in random grouped classes.

Let us now summarize findings in the specific problem areas of the two SRA inventories, and see what conclusions may be drawn. These conclusions are better supported for junior high school pupils than for elementary school pupils, because of the availability of comparative data from three samples at the former level.

Ability grouped junior high school pupils were found to report fewer problems on the "Looking Ahead" scale of the SRA Youth Inventory than comparable pupils in random grouped classes. These differences are most pronounced between superior boys and girls and average boys. It may be concluded that junior high school pupils in ability grouped classes have fewer problems relating to choosing a career, going to college, and finding a job than do comparable pupils in random grouped classes.

The "Getting Along With Others" scale differs somewhat on the two inventories. The results showed two significant between-treatment differences favoring District R pupils on the elementary school data. These results lead to the tentative conclusion that average pupils in the ability grouped classes at the elementary level have more social adjustment problems than comparable random grouped pupils. Three significant differences plus four differences approaching significance were found on the junior high school data all indicating higher problem scores for District R pupils. At this level, it may be concluded that there is a tendency for random grouped pupils at all ability levels to have a greater number of problems relating to social skills than comparable pupils in ability grouped classrooms.

Comparisons between fifth grade pupils at different ability levels showed a consistent trend for pupils of higher ability to list fewer social adjustment problems. No such tendency was apparent in the within-treatment comparisons at the junior high school level. It may be concluded tentatively that elementary school pupils of high ability report fewer social adjustment problems, regardless of the grouping treatment.

Comparisons between the treatments on the "Boy Meets Girl" scale showed eight statistically significant differences and one difference approaching significance. These differences all indicated a larger number of problems for District R junior high school pupils. From these data, it may be concluded that there is a consistent tendency for junior high school pupils in random grouped classes to report greater numbers of problems relating to dating, sex information, and sex mores than comparable pupils in ability grouped classes. These differences were most marked between superior pupils and least pronounced between slow pupils in the two treatments. Within-treatment comparisons on the "Boy Meets Girl" scale indicated that there was no relationship between ability level and problems in this area in either treatment.

Results on the "Basic Difficulty Key" of the Youth Inventory revealed one significant difference favoring District A, and five differences approaching significance in the same direction. These data suggest that if any significant between-treatment difference in pupil adjustment was present, this difference favored the ability grouping system. Because of the small differences found between the districts on the "Basic Difficulty Key," however, it seems safer to conclude that neither treatment is consistently associated with more indications of pupil maladjustment.

## REFERENCES

Remmers, H. H., & Bauernfeind, R. H. Manual for the SRA Junior Inventory, Form S. Chicago: Science Research Associates, 1955.

Remmers, H. H., Shimberg, B., & Drucker, A. J. Technical Supplement for the SRA Youth Inventory, Form A. Chicago: Science Research Associates, 1953.

# Chapter 10

# SELF-CONCEPT

## INTRODUCTION

The self-concept or phenomenal self as defined in contemporary psychology refers to the individual's perception of himself. That is, expressed in simple terms: the ideas, perceptions, and beliefs that form the image of himself that the individual has created. This picture that he has of himself may be highly realistic in the well adjusted individual, or it can be completely detached from reality as in the case of the psychotic who believes himself to be Napoleon. Regardless of the individual's self-concept, however, he tends to behave in a manner that he considers to be consistent with this concept. The self-concept is considered to be an important factor underlying the individual's behavior by most contemporary personality theorist. Wylie in her review of the self-concept points out that "...all the theories of personality which have been put forth within the last two decades assign importance to a phenomenal and/or nonphenomenal self-concept with cognitive and motivational attributes" (Wylie, 1961, p. 2).

Among theorists who assign an important role to the self-concept are included Adler (1924), Freud (1950), Fromm (1939), Horney (1937), Lecky (1945), Maslow (1954), Rogers (1951), and Snygg and Combs (1949). Some personality theorists including Rogers, Snygg and Combs, and Lecky, stress the self-concept as the principal determiner of behavior. Generally, these theorists consider that the need to preserve and enhance the self-concept is a basic human motivation.

Lecky (1945) stresses preservation of the self-concept and relates this variable to school behavior. He believes that the child who sees himself as weak, bad, or stupid will behave in ways that will tend to perpetuate this picture and refuse to accept the idea that he is good, strong, or clever. Jersild (1952) reflects a similar opinion concerning learning when he points out that "When a person resists learning that may be beneficial to him, he is, in effect, trying to protect or to shield an unhealthy condition. But, more broadly speaking, he is not actually protecting something unhealthy as such; he is trying to safeguard his picture of himself, his self-concept, the illusions concerning himself which he has built and which give him much trouble" (p. 114).

When we view the self-concept as a determiner of behavior and as a picture that the individual will maintain at considerable cost but is under some pressure to change when his behavior is contradictory to it, we can see some definite implications of this theory to the school situation and more specifically to the possible effects of ability grouping. Part of the picture that makes up the pupil's self-concept is concerned with the school. This picture, of course, is highly complex and involves such matters as the individual's estimate of his achievement, his perception of his status relative to other pupils in the class, his satisfaction with his school role, and his attitudes about study. Each aspect of the pupil's self-perception has a great many subtle shadings and is colored by his value system, the pressures placed upon him by the home and a great many other variables.

The two treatments employed in this research, ability grouping and random grouping, might be expected to influence the individual's self-concept in a number of ways. For example, a pupil who has been unaware of possessing superior ability during the years he has spent in a random grouped system may, upon being assigned to a superior class in an ability grouped system, become more actively aware of his academic status and incorporate this into his self-concept. The pupil of low ability may similarly be successful in a heterogeneous class in maintaining a higher self-perception of his ability than the objective reality of the situation would justify. Upon being placed in an ability grouped system, however, in which this unrealistic self-concept is directly contrary to his placement in a low ability group, he may find it no longer possible to maintain his previous estimate of his own ability and may thus find it necessary to revise downward some aspects of his self-concept.

On the other hand, the increased competition in the ability grouped class may tend to lower the self-concept of the bright pupil while the decreased competition faced by the slow pupil may raise his self-concept. Another circumstance that could lead to revision of the individual's self-concept in an ability grouped situation arises when his initial assignment to a section at one ability level is followed by subsequent reassignment to a lower or higher section.

The relative impact of the two treatments upon the individual's self-concept appears to be an important factor in the overall evaluation of ability grouping as compared to random grouping. Although it is possible on the basis of self-concept theory to make some tentative predictions concerning the effect of certain aspects of these two treatments upon the individual's self-concept, the complexity of the treatments make it difficult to predict their net effect. These predictions must, therefore, be tested with research evidence in order to provide us with any meaningful data on the overall effect of the two grouping systems on the self-concept. It is difficult, for example, to predict on the basis of self-concept theory whether the experiences of the low ability pupil in a heterogeneous class where it is possible for him to make direct comparisons between his performance and that of more able classmates would be more or less likely to lower his self-concept than his assignment to a low ability class in an ability grouped system where he would not be required to compete with pupils superior to himself but would be aware of his position in a low ability group and may well feel the stigma attached to membership in this group.

## WORKING HYPOTHESES

The purpose of including the Bills' Index of Adjustment and Values (IAV) and the other self-perception measures employed in this research was to test the null hypothesis that no significant changes in self-perception would occur between superior, average and below average pupils in the two grouping systems. The Index of Adjustment and Values, as employed in this research, yields four scores. These scores are: (1) Concept of Self, (2) Acceptance of Self, (3) Ideal Self and (4) Discrepancy Score, the latter obtained by subtracting Concept of Self from Ideal Self (Bills, undated).

A number of working hypotheses were developed with regard to the self-concept variables as they may operate in the ability grouping and random grouping systems.

### Between-Treatment Differences

It was hypothesized that, in comparisons between pupils of the same ability level in the two districts, two major treatment differences would operate that could have an effect upon the self-concept scores of the subjects. These variables are: (1) the awareness of the pupil in District A of the level at which he is placed, and (2) the necessity in District R for the pupil to compete with pupils of all ability levels.

Applying these variables to superior District A pupils, we can assume that they would, in all probability, be aware of their placement in superior classes and this placement should tend to develop higher self-concepts in the individual. On the other hand, placement in the superior section brings with it more intense competition and less of a feeling of being outstanding than the pupil may have experienced prior to the adoption of ability grouping in District A. This second factor, it was hypothesized, would lead District A superior pupils to develop higher Discrepancy Scores and lower Acceptance of Self scores. In other words, the bright pupil in the ability grouped classroom would be aware of his superiority but would also be aware that the other pupils in his group were equally or perhaps more superior and this awareness may lead him to be less self satisfied than he would be in a random grouped class.

It was hypothesized that average pupils would generally obtain lower scores in Self-Concept and Acceptance of Self and higher Discrepancy Scores in the ability grouped classroom because the individual would be more actively aware of his average status and less likely to overestimate his ability than would be the case in District R. Thus, District A average pupils would be somewhat obliged to face reality, but would be in no better competive position than if they were in a heterogeneous classroom.

It was considered that the same factors hypothesized to influence the self-concept of superior pupils would act somewhat differently upon the slow pupil. Identification with the slow group would tend to lower the pupil's self-concept while membership in the homogeneous class would make him better able to compete and this better competitive position might tend to enhance his self-concept. It was further reasoned, however, that the stigma of being identified with the slow group would outweigh the advantages gained by the individual's better competitive position and it was thus hypothesized that slow pupils in District A would have lower Self-Concept scores, lower Acceptance of Self scores and higher Discrepancy Scores than would slow pupils in District R. It will be noted that no working hypotheses were developed for the Ideal Self score, as this variable was not expected to change as a result of the grouping situation.

### Within-Treatment Differences

Within District A, it was hypothesized that superior pupils would have the higher Self-Concept scores, but would also have the higher Discrepancy Scores as compared with average and slow pupils. In District R, it was hypothesized that superior pupils would be higher than average and slow pupils on all four of the self-concept variables measured by the IAV.

It will be seen from the working hypotheses that we are concerned both with differences between the two grouping systems, which will be analyzed by comparing pupils of similar ability in the districts; and with differences within each system, which will be analyzed by comparing pupils of different ability levels. Although less directly pertinent to the goals of this project, we will also analyze sex differences occurring on the self-concept measures. These comparisons will involve boys and girls in the same grouping treatment and at the same ability level and it is hoped will give us greater insights into the nature and extent of sex differences in this important area.

We will also explore the question of what happens to the pupil in the ability grouped system when he is reassigned to a class at another ability level classification. It was hypothesized that pupils moving from a higher to a lower ability classification would obtain lower scores in Concept of Self and Acceptance of Self and higher Discrepancy Scores while pupils moving from a lower to a higher ability level classification would receive more favorable scores on these variables.

Although some data relating to self-concept were collected on all five of the research samples, more data were obtained from Sample IV and Sample VI than from the secondary samples.

## PREVIOUS RESEARCH RELATED TO SELF-CONCEPT AND GROUPING

### Self-Concept and Ability Grouping

Only four studies have been reported in the literature which are concerned with the relationship between ability grouping and self-concept. The first such study was carried out by Mann (1960) with a sample of 102 fifth grade pupils who had been ability grouped into four ability levels since entrance into the first grade. Grouping was based upon the results of group intelligence tests and reading readiness tests. Sections were labeled only by the teachers' names but teachers informally referred to them as "highest," "lowest," "second highest," and "second lowest" groups. Self-concept data used in this study were based on the subjects' anonymous responses to two questions. One of these was, "Tell me which fifth grade you are in," and the other was, "Tell me how you happen to be in this particular fifth grade group rather than some other group." About two-thirds of the lowest and highest groups identified their groups by ability level. A great majority of the two center groups identified their groups by the teacher's name. In response to being asked why they were in a particular group, 25 of the 30 pupils in the highest group answered in positive terms relative to ability or achievement such as "I'm smart," "I make good grades," "We work a little faster," and so on. In the lowest group, 14 of the 18 pupils answered in negative terms relative to ability or achievement such as "I am too dumb," "I am not so smart," "We don't know too much."

Mann's study clearly indicates that pupils in an ability grouped situation are more aware of the nature of their situation, but presents no direct evidence on the effect of grouping on the student. No control group was employed. It is possible that use of a control group would have revealed that low and high ability pupils in random groups have similar low and high perceptions of their ability and achievement. Furthermore, a questionnaire based on only two items would be expected to have rather low reliability. Mann's study would have been improved considerably by the use of a more extensive measure of self-concept. Other studies such as the work of Luchins and Luchins (1948), have also suggested that pupils in an ability grouped situation are aware of their grouping status regardless of efforts made by the school to conceal the grouping level. This awareness, coupled with the associations that the pupil makes with his ability level as indicated by Mann, certainly raise the question as to whether ability grouping may have undesirable effects upon the pupil's self-concept.

Drews' research (1962) with ability grouped and random grouped classes in ninth grade English was discussed in Chapter 4. This study also included measures of self-concept. Drews hypothesized that slow pupils would see themselves as relatively more competent in a homogeneous situation while the reverse would be true of superior pupils. Two self-concept measures were employed to test this hypothesis. The Ability Self-Concept Rating consisted of a single question asking the student to compare his ability with that of his classmates and rate himself as above average, about average, or below average. The other measure used was the 20 item Concept-of Self-as-a-Learner Scale which Drews developed from Bills' Index of Adjustment and Values.

The Ability Self-Concept Rating was administered at the beginning and end of the experimental year. Mean scores were analyzed using the t-test to determine if significant differences existed between homogeneous and heterogeneous classes. Results on this single item scale are summarized in Table 10-1. Only one significant difference was found on the pretest analysis--favoring slow pupils in the homogeneous sample. At the end of the one year experiment, between-treatment comparisons revealed that superior pupils in heterogeneous groups and slow pupils in homogeneous groups made significantly higher scores on this measure.

Drews administered the Concept-of-Self-as-a-School-Learner Scale at the end of the study only. Table 10-2 shows her between-treatment results. It will be noted that although the heterogeneously grouped superior pupils obtained higher mean scores, the differences were not significant. Only one significant between-treatment difference occurred, favoring the slow boys in the homogeneous groups.

Drews also made comparisons within each of her treatments between pupils of different ability levels. All six comparisons made between superior and average pupils showed statistically significant differences favoring the superior groups. Differences between average and slow pupils showed somewhat smaller differences with three of the six differences reaching statistical significance. These data are summarized in Table 10-3.

An important study by Goldberg and her associates has not been completed at the time of this writing. She reports, however, that self-attitudes were significantly affected by grouping.

> In narrow range groups the gifted children showed a significant decrease in their self-appraisals which were considerably inflated to begin with. On the other hand, the slow children in narrow-range groups showed increased self-appraisals which brought them more nearly to the level of the average. In the broad range groups the reverse phenomenon was apparent. The gifted showed considerable increases in self-appraisal while the slow decreased significantly. Depending upon your point of view, effects of grouping on self-attitudes may be perceived as desirable or undesirable. The effects of grouping were even more startling with regard to self-aspirations. There, the broad range group, and especially the presence of gifted children tended to act as a damper on the aspirational ceilings of the slower pupils.[1]

---

[1] Quoted from a personal communication from Dr. Goldberg, dated September 27, 1963, with author's permission.

## Table 10-1

Comparison of Mean Ability Self-Concept Ratings by Achievement Group and Sex in Homogeneous and Heterogeneous Classes[c]
(From Drews, 1962, p. 5-8)

| Class | Superior | | | Average | | | Slow | | |
|---|---|---|---|---|---|---|---|---|---|
| | Boys | Girls | Total | Boys | Girls | Total | Boys | Girls | Total |
| **Pre-Testing** | | | | | | | | | |
| Homogeneous | 2.46 | 2.67 | 2.61 | 2.11 | 2.02 | 2.06 | 1.95 | 2.05 | 1.98 |
| Heterogeneous | 2.33 | 2.45 | 2.39 | 2.12 | 2.16 | 2.14 | 1.62 | 1.88 | 1.71 |
| *t* | .70 | 1.56 | 1.78 | 1.41 | 2.14 | 1.53 | 1.79 | 1.40 | 2.13[a] |
| **Post-Testing** | | | | | | | | | |
| Homogeneous | 2.12 | 2.18 | 2.19 | 2.19 | 2.10 | 2.15 | 2.21 | 2.30 | 2.24 |
| Heterogeneous | 2.92 | 2.73 | 2.83 | 2.15 | 2.19 | 2.17 | 1.69 | 1.75 | 1.71 |
| *t* | .10 | 3.76 | 6.43[b] | .40 | .95 | .36 | 2.32 | 2.81 | 3.97[b] |

[a] $p < .05$

[b] $p < .01$

[c] Scale for Ability Self-Concept Rating:
3 = above average
2 = about average
1 = below average

## Table 1[illegible]-2

### Comparison of Ratings of Concept-of-Self-as-School-Learner by Achievement Group and Sex and Between Homogeneous and Heterogeneous Classes[c]

(From Drews, 1[illegible]62, p. 5-11)

| Group and Sex | Homogeneous | t | Heterogeneous |
|---|---|---|---|
| Superior | | | |
| Boys | 17.3 | .28 | 18.3 |
| Girls | 21.0 | 1.71 | 22.7 |
| Total | 19.9 | .43 | 20.4 |
| Average | | | |
| Boys | 13.6 | 1.50 | 11.9 |
| Girls | 14.6 | .37 | 13.4 |
| Total | 14.1 | 1.46 | 12.7 |
| Slow | | | |
| Boys | 11.2 | 6.11[b] | 4.2 |
| Girls | 10.8 | .17 | 7.7 |
| Total | 11.1 | 2.20[a] | 5.5 |

[a] $p < .05$

[b] $p < .01$

[c] Concept-of-Self-as-School-Learner scale consisted of twenty questions with four-place response scale ranging from "just (or) mostly like me" (3) to "not like me" (1).

A portion of the California study of gifted children (Simpson and Martinson, 1961) showed that eighth grade pupils in special classes made greater gains on Self Acceptance as measured by the California Psychological Inventory than did pupils of comparable I.Q. who were assigned to heterogeneous control classes. These findings agree with those reported by Goldberg but disagree with Drews (1962). Perhaps these findings differ because Drews' study made special curriculum adjustments for the heterogeneous groups while the other two studies did not.

## Other Pertinent Research

A number of other studies dealing with self-concept, although not dealing directly with the ability grouping situation, contain evidence that may give the reader some insights into the self-concept and how it operates in the educational situation. Perhaps the most pertinent of these studies are concerned with the relationship between self-concept and achievement. Conklin (1940), in a case study aimed at discovering differences between academically successful and academically unsuccessful gifted children, found a number of variables in the histories of unsuccessful children that would suggest that these children had low self-concepts.

Perkins' research (1958) employed a Q-sort instrument for measuring self-concept. Pupils completed two sorts; one for self-concept, and one for ideal-self. Self-ideal-self congruency was determined by correlating the two Q-sorts. Subjects in this research consisted of 251 pupils from four fourth grade and four sixth grade classrooms. The reliability of the Q-sort for the entire sample was .65. Self-ideal-self congruency is similar to the Discrepancy Score employed by Bills in the IAV although arrived at in a somewhat different manner. Perkins compared self-ideal-self congruency with a number of other variables. Among the significant differences pertinent to this report was the finding that the self-ideal-self congruency scores of girls were significantly greater than those of boys. This would suggest that a smaller discrepancy between self-concept and ideal-self for girls might be found on the IAV. Perkins also found that the correlation between self and ideal-self increased between the fourth and sixth grade.

Both the Q-sort instrument and the California Achievement Test were administered to the sixth grade pupils before and after the study, a period of about six months. Product-moment correlations were computed between changes in self-ideal-self congruency and changes in academic achievement. These correlations ranged from -.06 to .16 for the four classes. None of the correlations were significant, thus suggesting that changes in self-ideal-self congruency are not related to changes in school achievement.

This research also sought to determine the relationship between changes in self-ideal-self congruency and peer group acceptance. Peer group acceptance was measured using a sociometric questionnaire in which pupils listed their "very best friends." The child's peer acceptance score was the number of times he was listed on the sociometric instrument. Product-moment correlations between changes in self-ideal-self congruency, and changes in peer acceptance were .04 for the fourth grade sample and .03 for the sixth grade sample, thus indicating no relationship.

Perhaps the major weakness in Perkins study was the relatively short period over which his data were collected--less than six months. It seems quite possible that changes in self-ideal-self congruency may lag considerably behind changes in school achievement and peer status and thus not be adequately reflected in such a short period.

Table 10-3

Comparison of Ratings of Concept-of-Self-as-School-Learner Between Achievement Groups in Homogeneous and Heterogeneous Classes[c]
(From Drews, 1962, p. 5-11)

| Class and Sex | Superior | t | Average | t | Slow |
|---|---|---|---|---|---|
| Homogeneous | | | | | |
| Boys | 17.3 | 2.45[a] | 13.6 | .96 | 11.2 |
| Girls | 21.0 | 4.39[b] | 14.6 | 1.78 | 10.8 |
| Total | 19.9 | 3.89[b] | 14.1 | 2.14[a] | 11.1 |
| Heterogeneous | | | | | |
| Boys | 18.3 | 2.64[b] | 11.9 | 2.74[a] | 4.2 |
| Girls | 22.7 | 3.30[b] | 13.4 | 1.69 | 7.7 |
| Total | 20.4 | 3.89[b] | 12.7 | 3.38[b] | 5.5 |

[a] $p < .05$

[b] $p < .01$

[c] Concept-of-Self-as-School-Learner scale consisted of twenty questions with four-place response scale ranging from "just (or) mostly like me" (3) to "not like me" (1).

Benjamins (1950) has reported an interesting study on the relationship between changes in the pupil's self-concept concerning his intellectual level and changes in his performance on a group intelligence test. Benjamins' study was carried out to test Raimy's theory that the self-concept is a focal point in determining much of the individual's behavior and specifically to test the theoretical construct that regards the self-concept as both influencing and being influenced by the individual's behavior. In this research 48 subjects from two high school classroom groups were studied. Pupils were first asked to indicate what they believed their rank to be within the group on popularity, grades, intelligence, and other variables. In addition they were asked to give the upper and lower limits of the range of ranks into which they considered their ranks would surely fall. Immediately after these first self-rankings, form A of the Otis Self-Administering Test of Mental Ability was administered. On the following day subjects were presented individual reports giving what was purported to be their ranks on the test they had taken on the previous day. The ranks reported to the students were systematically altered, being either one rank higher or one rank lower than the limits in which the subject had indicated he surely would fall. The subjects were then asked to re-rank themselves. The second self-rankings were to represent the pupil's estimate of what his real rank in intelligence would be if the whole truth were known. It was explained that the tests were not perfect and did not tell the whole truth about the pupil's intellectual level. Immediately after this and without previous warning, form B of the Otis test was administered and followed by a questionnaire that collected more information concerning the individual's self-perception of his ability. The investigator then made predictions of each individual's score of the second intelligence test, based on the false report and the pupil's answers to the questionnaire regarding his reactions to the report. Analysis of the relationship between the expected and observed changes in the intelligence test scores yielded a chi-square significant beyond the .01 level. Results of this study, therefore indicated that changes in the individual's self-concept as induced by the false reports of his first test results led to similar changes in his performance on a latter test, thus reflecting the individual's effort to achieve a balance between his performance and his self-concept.

This is a rather well controlled experimental study that demonstrates that the changes in the individual's self-concept can lead to changes in his behavior, a position taken by several of the theorists in this area including Raimy (1948), Lecky (1945), and Jersild (1952), among others. Perhaps the major weakness of Benjamins' study is the extremely short time over which data were collected. It would be interesting to learn whether the changes in self-concept induced by the false report of the individual's intelligence test score would have persisted for any length of time and would have led to a lowering of performance on a second test administered after a longer interval. It will be recalled that the two estimates made by the individual of his rank on the two intelligence tests, and the final questionnaire were all administered within a period of two days.

One aspect of the self-concept theory indicates that when faced with a need to change behavior, this need may or may not be perceived as a threat to the individual's self-concept. If the change is perceived as a threat, the individual will resist the change and not alter his behavior, but if the change is not perceived as a threat the individual will make the change and also make the necessary alterations in his self-concept to bring it back into line with his changed behavior. A study by Roth (1959) was designed to test this theoretical construct. Freshman subjects were drawn from three reading improvement classes at the University of Texas. These were voluntary non-credit classes requiring only that the student read at a minimum of 250 words per minute with 75 per cent comprehension. The 54 subjects selected were administered a diagnostic reading test, a sentence completion technique, and two Q-sorts indicating self-concept and ideal-self. Students were then exposed to 14 sessions of reading improvement program after which the initial measures were again administered. These measures yielded information on reading improvement, changes in self and ideal-self, and the amount of defensiveness of the individual concerning his concept of self in a number of different contexts. Pupils were sorted into three groups, "improvers," "non-improvers," and "attritions" on the basis of their second score on the diagnostic reading test. It was hypothesized that there would be significant differences in the amount of defensiveness among the three reading groups with the attrition group being the most defensive followed by the non-improver, and improver groups. This hypothesis was supported by the data, indicating that the more defensive the individual becomes of his self-concept the less likely he is to improve or change his behavior. This pattern held for general defensiveness, defensiveness in the self-concept and defensiveness in the individual's perception of himself as a reader. The data from this study generally support the theoretical construct concerning relationships between self-concept and behavior changes and demonstrate the importance of the individual's self-concept and the degree in which he is defensive about this concept as a factor in his school achievement.

A study by Bruck (1959) was concerned with relationship between self-concept and achievement of 300 pupils in grades three, six, and eleven. The self-concept was measured using the Draw-a-Person test, while grade point average was used to measure academic achievement. This study found that self-concept and grade point average were significantly correlated at all three grade levels. Significant age differences were found between pupils at the third and eleventh grade levels in the amount of relationship between self-concept and grades. Significant sex differences were also found at the third and eleventh grade levels. From the standpoint of the Utah study, the most pertinent information obtained from Bruck's research is that self-concept and grades are related. Although Bruck does not establish a cause and effect relationship between these two variables, his work at least opens the possibility that such a relationship does exist.

Checkering (1958) also studied the relationship between achievement and self-concept. The self-concept of 109 ninth grade students was measured using the Q-sort technique, and self-ideal-self discrepancy scores were compared with achievement. A significant inverse relationship was found between achievement and the discrepancy score, indicating that higher achieving students obtained smaller discrepancies between self and ideal-self thus implying better adjustment. Among the other findings in this research, Checkering found that the ideal-self measure was of little value in itself. This finding formed the basis for establishing null hypotheses on the ideal-self score in the Utah study.

Although rather limited, the research in this area seems to indicate rather clearly that school behavior and the self-concept are related, with each variable probably causing changes in the other under certain circumstances. The work of Drews (1962), Mann (1960), and Goldberg all contain evidence to indicate that ability grouping can alter self-concept.

## PROCEDURE

### Sample and Measures

### The Index of Adjustment and Values

The principal measure of self-concept employed in this research was the Index of Adjustment and Value, developed by Robert E. Bills (Bills, et. al., 1951). The initial form of this measure was published in 1951 and a considerable body of research and evaluation has been carried out on the initial form and subsequent forms since that time. The Index of Adjustment and Values is currently available in four forms: (1) the elementary school form to be used in grades 3, 4, and 5; (2) the junior high school form for grades 6, 7, and 8; (3) the high school form for grades 9, 10, 11; and (4) the original adult form which is also used in grade 12. The first three of these forms were employed in this research project. The original form of the IAV was developed by selecting 124 trait words which the authors believed presented clear examples of self-concept definitions, from Allport and Odbert's list of 17,953 trait words. In developing the IAV this list was reduced to 49 words that are the basis of the adult form. Trait words for the other three forms were developed by interviewing approximately 850 children in grades one to eleven, and compiling word lists related to self-concept from these interviews. From these lists a number of trial forms were developed and revised leading ultimately to the forms used in this research. The format of the test and administration instructions were also changed by Bills to make them more understandable for the pupils at the lower grade levels.

The high school form contains 37 trait words, the junior high school form contains 35 words, and the elementary school form contains 19. In all of these forms the pupil answers three questions about himself: he rates himself with respect to the trait word (concept-of-self), tells how he feels about being this sort of person (acceptance-of-self), and how he would like to be with respect to each trait (ideal-self). On the high school form each of the ratings is made on a five-point scale, the subject completing all three ratings for each trait before going on to the next. On the junior high school form, instead of using a five-point rating system, the pupils make checkmarks on a three-point scale. On the elementary school form, a three-point scale is also used, but the pupil's task is further simplified in that the administrator reads three questions to the pupil for each trait word, and the pupil is required only to circle the appropriate responses on his answer sheet. Pre-testing of the junior high school and elementary forms of the IAV in the Utah study indicated that more detailed instructions and more examples were necessary if most children were to fully understand what they were to do. Therefore, although Bills' instructions were followed in essence, additional examples were given and proctors were trained to check each pupil during the administration of the measure and assist those who appeared not to understand the instructions.

In addition to obtaining "self" responses on each of the trait words, the IAV also requires the pupil to make similar responses concerning "others" in his group. In this second part of the test the pupil is asked to fill out the form as he believes the average member of his group would fill it out for himself. The "others" portions of the IAV were not employed in the research reported herein.

## Intercorrelational Data

In order to further assist the reader in interpreting the self-concept data, intercorrelational matrixes were calculated between the four scores obtained on the elementary school IAV administered at the fifth grade level and the junior high school IAV, administered at the sixth and seventh grade levels. These correlations were based on the Sample IV data and may be found in Table 10-4. Also included in this table is the intercorrelational matrix for the adult IAV given in the IAV test manual.[1] The manual does not give intercorrelational matrixes for the elementary or junior high school forms of the test. It may be seen that the correlations for the adult form differ somewhat from those obtained in this research for the elementary and junior high school forms. The correlation between Concept of Self and Acceptance of Self is lower in both of these forms than the figure given by Bills. The negative correlation between Concept of Self and Discrepancy Score is slightly lower but generally in agreement with Bills' data.[2] The correlation between Acceptance of Self and the Discrepancy Score is also considerably lower than that given by Bills. While Bills gives a low negative correlation between the Ideal Self and the Discrepancy Score, it will be noted in Table 10-4 that moderate positive correlations were found between these two variables on the elementary and junior high school forms. It appears likely that these lower correlations are at least partially due to the lower reliability that one would expect on the elementary and junior high school forms. The moderate positive relationship between Ideal Self and Discrepancy Score seems noteworthy. It should be recalled that the Discrepancy Score as calculated in the Utah study is the difference between the Concept of Self and Ideal Self score, and therefore a portion of the correlation between these scores would be spurious.

## Reliability

Reliability coefficients reported by Bills for the junior high school and elementary school Index of Adjustment and Values are relatively high. Bills reports split-half reliabilities on the Self Concept score for different samples ranging from .74 to .93. For Acceptance of Self reliabilities ranging from .80 to .92 are reported, and for Ideal Self ranging from .85 to .94. Test-retest reliabilities over a period of six weeks are reported to range from .61 to .71 for Concept of Self, .74 to .89 for Acceptance of Self and .53 to .81 for Ideal Self.[3] Bills reports corrected split-half reliabilities for one sample of 80 pupils on the elementary school IAV as .48 for Self Concept, .69 for Acceptance of Self and .72 for Ideal Self.

A total of 455 pupils in District R Sample IV completed all three of the administrations of the IAV scheduled for this sample. In order to get some further insight into the relationship between the elementary and junior high school forms of the

---

[1] Bills, Robert E. Manual, Index of Adjustment and Values, (undated), p. 62.

[2] The Discrepancy Score was calculated by a different method in the Utah study than recommended by Bills.

[3] Ibid., p. 13-14.

Table 10-4

Intercorrelational Matrices Between IAV Scores[a]

| | Elementary School IAV (5th Grade) | | | Junior High School IAV (6th Grade) | | | Junior High School IAV (7th Grade) | | |
|---|---|---|---|---|---|---|---|---|---|
| | District A - N=422 | | | District A - N=386 | | | District A - N=349 | | |
| Variable | 2 | 3 | 4 | 2 | 3 | 4 | 2 | 3 | 4 |
| 1. Concept of Self | .32 | .15 | -.75 | .60 | .31 | -.61 | .54 | .34 | -.65 |
| 2. Acceptance of Self | | .16 | -.21 | | .28 | -.31 | | .33 | -.25 |
| 3. Ideal Self | | | .41 | | | .48 | | | .45 |
| 4. Discrepancy Score | | | | | | | | | |

| | Elementary School IAV (5th Grade) | | | Junior High School IAV (6th Grade) | | | Junior High School IAV (7th Grade) | | | Adult IAV Data from Manual[b] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | District R - N=580 | | | District R - N=621 | | | District R - N=570 | | | College - N=300 | | |
| | 2 | 3 | 4 | 2 | 3 | 4 | 2 | 3 | 4 | 2 | 3 | 4 |
| 1. Concept of Self | .38 | .32 | -.81 | .67 | .42 | -.63 | .52 | .42 | -.70 | .90 | .55 | -.83 |
| 2. Acceptance of Self | | .15 | -.29 | | .30 | -.46 | | .26 | -.37 | | .34 | -.67 |
| 3. Ideal Self | | | .23 | | | .34 | | | .31 | | | -.15 |
| 4. Discrepancy Score | | | | | | | | | | | | |

---

[a] Based on Sample IV data.
[b] Bills, Robert E. Manual, Index of Adjustment and Values, (undated), p. 62.

measure and some estimates of the stability of *IAV* scores, the four *IAV* subscores earned by these pupils on the three administrations of the measure were correlated. The results may be found in Table 10-5. It will be noted that the correlations between the elementary school *IAV* administered at the fifth grade level and the junior high school *IAV* administered at the sixth and seventh grade levels were quite low, ranging from .09 between the fifth and seventh grade *Ideal Self* scores to .28 between the fifth and sixth grade *Concept of Self* scores. On the other hand, correlations between the sixth and seventh grade administrations of the junior high school *IAV* ranged from .41 to .51. These data seem to indicate that the elementary and junior high school levels of the test may be measuring variables that are only roughly comparable. The correlations between the sixth and seventh grade administrations seem to indicate that the junior high school form of the test has moderately good long term stability.

## *Sense of Personal Worth*

In addition to the *Index of Adjustment and Values*, a number of other measures related to self-concept were administered. One of these was the *Sense of Personal Worth* scale taken from the *California Test of Personality*. This measure was administered to Sample IV during their fifth grade year. The reliability coefficient for this subtest is reported to be .79 in the test manual (Thorpe, *et*. *al*., 1953). This measure will be discussed in Chapter 12.

## *Self-Acceptance*

The purpose of the *Self-Acceptance Scale* of the *California Psychological Inventory* as given in the manual is "to assess factors such as sense of personal worth, self-acceptance, and capacity for independent thinking and action" (Gough, 1957, p.12). This measure was administered to Sample VII pupils during their ninth grade year, and to Sample VI pupils during their eighth grade year. The test-retest reliability coefficient of this scale based on a sample of high school students is reported in the test manual as .71 and .72 for male and female students respectively. Two studies are reported by Gough indicating significant correlations between this scale and ratings of self-acceptance, thus providing some evidence of concurrent validity. In one of these studies, biserial correlations of .46 between the scores of high school boys and principals' ratings of self-acceptance and .39 between girls' scores and principals' ratings are reported. A somewhat more detailed discussion of the *California Psychological Inventory* may be found in Chapter 12.

## *About Myself Scale*

Both the *SRA Youth Inventory*, *Form A*, and the *SRA Junior Inventory*, *Form S*, contain scales in which the individual indicates problems relating to himself. Pupils in Samples IV, VI, VII, and VIII were administered one of these two problem check lists. Reliability coefficients for this scale in the *Youth Inventory* are recorded in the manual as .88, .90, and .73, based upon three different samples, (Remmers, *et*. *al*., 1953). The reliability coefficient for the parallel scale in the *Junior Inventory*, is given in the manual as .81 on a sampling of 3,000 children (Remmers and Bauernfeind, 1957). A more detailed discussion of these measures will be found in Chapter 11.

Table 10-5

Correlations Between IAV Scores Administered
in 5th, 6th and 7th Grades to Sample IV,
District R (N = 455)

| | Concept of Self | | Acceptance of Self | | Ideal Self | | Discrepancy Score | |
|---|---|---|---|---|---|---|---|---|
| | 6th | 7th | 6th | 7th | 6th | 7th | 6th | 7th |
| 5th Grade | .28 | .27 | .25 | .20 | .15 | .09 | .16 | .18 |
| 6th Grade | | .51 | | .46 | | .46 | | .41 |

# RESULTS

## Sample IV

### Introduction

Sample IV was administered the Bills Index of Adjustment and Values at the fifth, sixth, and seventh grade levels. At the fifth grade level, the elementary form was employed while the junior high school form was administered at the sixth and seventh grade level. The Sample IV data provide us with our best information concerning the relationship at the elementary level between the treatments and the four scores obtained from the Index of Adjustment and Values.

### Sex Differences on Elementary IAV

In order to determine whether sex differences were present on the Elementary School Index of Adjustment and Values, boys were compared with girls at each ability level in both districts on the four IAV scores. It was found that there were no significant differences between boys and girls in either district on any of the four IAV variables. Means and standard deviations for the 12 subgroups may be found in Appendix E. The reader will note in comparing boys and girls at the same ability levels that differences were extremely small, in most cases falling far short of statistical significance. On the basis of these results, the remainder of the analysis dealing with the Elementary School IAV scores for Sample IV were carried out combining the sexes at each ability level.

### Classification by Ability

During the final year of the project, Sample IV pupils were in the seventh grade. At this level, pupils were classified in terms of their ability level in both mathematics and science. District A placed pupils in ability groups in these two subjects and equivalent ability level classifications were established for pupils in District R. Thus, although the vast majority of pupils in both districts were found to be in the same ability levels in both mathematics and science, there were some pupils in each district who were classified in a different level in these two subjects. It seems highly unlikely that this difference would have any effect upon self-concept in District R in view of the fact that pupils were placed in random grouped classes and were unaware of the ability levels that had been assigned them in the research project. In District A, however, it would seem possible (although unlikely) that placement of a pupil in a higher mathematics ability section might have a different effect upon self-concept than placement of a pupil in a higher science ability level. In order to check this possibility, Sample IV pupils were classified into those in the same ability level in both mathematics and science, those placed in a higher section in mathematics and those placed in a higher section in science. Analysis of variance was then used to determine whether significant differences existed in self-concept scores for these various subgroups. Three comparisons were made in each district on each of the four self-concept variables. Pupils at the same level in both subjects were compared with pupils who were higher in mathematics and pupils who were higher in science, and finally pupils who were higher in mathematics were compared with pupils who were higher in science. Of the 12 comparisons made for District R pupils, none attained statistical significance and in most cases, the mean scores of the three groups compared were within a point or so of each other. In District A, one of these differences reached statistical significance. On the Ideal Self score, pupils who were in a higher ability grouped class in science were significantly higher at the .05 level than pupils who were in a higher ability grouped class in mathematics. In view of the lack of other differences that might indicate a real trend, it is the opinion of the investigator that this single difference was due to chance. On the basis of the general lack of significant differences in this analysis and the relatively small number of cases who were grouped at different ability levels in mathematics and science, it was decided that in the analysis of results obtained with variables other than achievement, secondary pupils in all samples would be classified into ability levels on the basis of a single subject--mathematics, rather than introduce the additional complexity involved in carrying out the analysis twice or setting up additional groups of pupils who were at different ability levels in the two subjects in which ability grouping was employed at the junior high and high school level.

## Sex Differences on Junior High School IAV

Differences between the mean scores of boys and girls from the same district and of comparable ability level on the Junior High School IAV are given in Table 10-6. A comparison of boys' and girls' groups on Concept of Self showed a fairly consistent trend for girls in this sample to receive higher scores than boys. Of the six comparisons for District A, three reached statistical significance, all favoring girls' groups. In District R, a similar trend was apparent, although somewhat less marked, with only one of the six sex differences reaching statistical significance.

Analysis of Acceptance of Self means, revealed only one statistically significant difference favoring District A superior girls. Although girls tended to have slightly higher Acceptance of Self scores in District A, this trend did not hold for District R where the largest mean differences obtained favored the boys' subgroups. A lack of any consistent pattern in these data suggests that the single significant difference that emerged was due to chance.

A comparison of the mean scores of boys and girls on Ideal Self, indicated a consistent tendency for girls to receive higher mean scores on this variable. In District A, five of the six differences reached statistical significance. The District R differences were somewhat smaller with only two of the six being statistically significant. Nonsignificant differences all followed this same trend.

Differences between the means of boys and girls on the Discrepancy Score were small and inconsistent. Only one statistically significant difference emerged in each of the districts. These differences were both found in the sixth grade data, favoring superior girls in District A and slow girls in District R. Both of these differences barely reached significance at the .05 level.

## Concept of Self Data

These data for Sample IV are summarized in Table 10-7. During the first year of the project, the mean scores of all Sample IV groups were very similar varying less than two points from the lowest to the highest group mean. One significant difference, however, was found between treatments. District R superior pupils scored significantly higher on Concept of Self than pupils of comparable ability in District A.

## Table 10-6

### Sex Differences on Junior High School IAV Scores for Sample IV

| | 1960-61 IAV | | | 1961-62 IAV | | |
|---|---|---|---|---|---|---|
| | 6th Grade | | | 7th Grade | | |
| Groups | $M_{Boys}$ | $M_{Girls}$ | t | $M_{Boys}$ | $M_{Girls}$ | t |
| **Concept of Self** | | | | | | |
| District A | | | | | | |
| Superior Boys vs. Superior Girls | 83.11 | 87.68 | 2.51 | 88.21 | 91.00 | N.S. |
| Average Boys vs. Average Girls | 82.48 | 85.41 | 2.06 | 84.88 | 87.69 | 2.31 |
| Slow Boys vs. Slow Girls | 84.72 | 86.61 | N.S. | 82.96 | 85.07 | N.S. |
| District R | | | | | | |
| Superior Boys vs. Superior Girls | 90.29 | 90.47 | N.S. | 93.32 | 93.12 | N.S. |
| Average Boys vs. Average Girls | 88.92 | 89.35 | N.S. | 91.92 | 92.63 | N.S. |
| Slow Boys vs. Slow Girls | 86.04 | 90.85 | 2.44 | 88.00 | 92.05 | A.S.[a] |
| **Acceptance of Self** | | | | | | |
| District A | | | | | | |
| Superior Boys vs. Superior Girls | 85.52 | 91.26 | 2.38 | 88.77 | 89.60 | N.S. |
| Average Boys vs. Average Girls | 87.64 | 89.91 | N.S. | 88.63 | 90.98 | N.S. |
| Slow Boys vs. Slow Girls | 86.26 | 88.89 | N.S. | 86.88 | 91.13 | N.S. |
| District R | | | | | | |
| Superior Boys vs. Superior Girls | 90.51 | 89.73 | N.S. | 94.13 | 91.82 | A.S. |
| Average Boys vs. Average Girls | 92.00 | 90.54 | N.S. | 95.00 | 92.62 | A.S. |
| Slow Boys vs. Slow Girls | 88.66 | 91.93 | N.S. | 92.55 | 94.26 | A.S. |
| **Ideal Self** | | | | | | |
| District A | | | | | | |
| Superior Boys vs. Superior Girls | 100.28 | 101.16 | N.S. | 101.86 | 103.15 | 2.08 |
| Average Boys vs. Average Girls | 91.92 | 95.31 | 1.98 | 95.31 | 100.83 | 4.88 |
| Slow Boys vs. Slow Girls | 88.38 | 94.94 | 2.32 | 90.19 | 97.20 | 2.18 |
| District R | | | | | | |
| Superior Boys vs. Superior Girls | 100.24 | 101.83 | 1.99 | 101.93 | 102.59 | N.S. |
| Average Boys vs. Average Girls | 97.97 | 99.24 | N.S. | 99.97 | 101.96 | 2.75 |
| Slow Boys vs. Slow Girls | 88.75 | 91.88 | N.S. | 92.77 | 95.98 | N.S. |

| Groups | 1960-61 IAV |  |  | 1961-62 IAV |  |  |
|---|---|---|---|---|---|---|
|  | 6th Grade |  |  | 7th Grade |  |  |
|  | $M_{Boys}$ | $M_{Girls}$ | t | $M_{Boys}$ | $M_{Girls}$ | t |
| Discrepancy Score |  |  |  |  |  |  |
| District A |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 17.53 | 14.05 | 2.04 | 13.98 | 12.15 | N.S. |
| Average Boys vs. Average Girls | 14.28 | 14.05 | N.S. | 13.94 | 15.41 | N.S. |
| Slow Boys vs. Slow Girls | 11.30 | 12.23 | N.S. | 13.80 | 17.00 | N.S. |
| District R |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 11.25 | 12.77 | N.S. | 9.39 | 10.06 | N.S. |
| Average Boys vs. Average Girls | 11.26 | 12.46 | N.S. | 10.28 | 10.74 | N.S. |
| Slow Boys vs. Slow Girls | 9.42 | 5.67 | 2.05 | 10.53 | 8.62 | N.S. |

[a] Approaching Significance - t = 1.65 - 1.95

Table 10-7

Sample IV - Differences in Concept of Self Scores Over a Three Year Period

| | 5th Grade Significant Differences | | | | 6th Grade Significant Differences | | | | 7th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N[a] | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 130 | 44.14 | | | 54 | 83.11 | | | 43 | 88.21 | | gr.3,5<br>1.99;2.38 |
| 2. Superior Girls | | | | | 57 | 87.68 | | | 40 | 91.00 | | gr.4,6<br>2.30,2.23 |
| 3. Average Boys | 212 | 44.36 | | | 100 | 82.48 | | | 104 | 84.88 | | |
| 4. Average Girls | | | | | 97 | 85.41 | | | 106 | 87.69 | | |
| 5. Slow Boys | 81 | 44.48 | | | 39 | 84.72 | | | 26 | 82.96 | | |
| 6. Slow Girls | | | | | 18 | 86.61 | | | 15 | 85.07 | | |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 208 | 45.80 | gr.1-2[b]<br>3.90 | gr.11-12<br>A.S. | 140 | 90.28 | gr.1<br>5.10 | gr.9,11<br>3.56;2.46 | 119 | 93.32 | gr.1<br>3.63 | gr.11<br>3.63 |
| 8. Superior Girls | | | | | 146 | 90.47 | gr.2<br>A.S. | | 111 | 93.12 | | |
| 9. Average Boys | 190 | 45.01 | | | 99 | 88.92 | gr.3<br>4.83 | | 116 | 91.92 | gr.3<br>5.86 | gr.11<br>2.48 |
| 10. Average Girls | | | | | 107 | 89.34 | gr.4<br>4.83 | | 114 | 92.63 | gr.4<br>4.33 | |
| 11. Slow Boys | 95 | 45.56 | | | 89 | 86.04 | | | 53 | 88.00 | gr.5<br>A.S. | |
| 12. Slow Girls | | | | | 41 | 90.85 | gr.6<br>A.S. | | 42 | 92.05 | gr.6<br>2.38 | |

[a] Groups not divided by sex for fifth grade data.

[b] Indicates group 7-8 is significantly higher than group 1-2 on the Concept of Self mean with a t of 3.90.

Comparisons between pupils at the average and slow levels in the two districts yielded no significant differences although the District R samples were slightly higher in both instances. No significant differences occurred in the within-treatment comparisons, that is, those comparisons between pupils of different ability levels within each district.

Thus, we find that exposure to ability grouping for a little less than one school year seemed to have had little influence upon the self-concept of the pupils involved when these scores were compared with the District R pupils who remained in random grouped classrooms.

Although following somewhat similar patterns, the Concept of Self scores obtained by Sample IV at the sixth grade level yielded a number of statistically significant differences. Again, it will be noted that District R pupils generally scored higher than District A pupils on this measure. In comparing groups of similar ability, we find District R pupils were significantly higher than their counterparts in District A in four of the six between-treatment comparisons. Only the difference between slow boys in the two districts failed to reach or approach statistical significance.

Comparisons indicated no within-treatment differences in District A. In District R, two statistically significant level differences occurred, superior boys having a higher mean than both average and slow boys. A pattern began to emerge in these data that suggests a relationship between ability level and Concept of Self scores within each district.

During the final year of the study, the patterns that began to emerge in the sixth grade data became clearer. Mean scores obtained during the seventh grade show District R pupils to be significantly higher than comparable pupils in District A in four of the six between-treatment comparisons. Within-treatment differences now reveal a consistent pattern in both districts, and for both sexes, showing the superior pupils to have the highest Concept of Self scores, the average pupils the next highest and the slow pupils, the lowest. In District A, four of the six level differences reached statistical significance; while two were significant in District R.

## Discussion

In reviewing the Concept of Self data for Sample IV, several important findings seem to emerge. First, the very slight differences existing both within and between the districts during the fifth grade became greater during the sixth grade and still greater at the seventh grade level. This trend could indicate a cumulative effect of the two treatments upon self-concept brought about by an increased awareness of the grouping. It could also suggest that pupils place a higher value upon their achievement status as they mature and move into the junior high school. It is interesting to note in Table 10-7 that the mean Concept of Self scores of all groups went up between the sixth grade and seventh grade testing, with the exception of District A slow pupils. This exception may lend support to the argument that ability grouping is destructive to the self-concept of the slow pupil as it tends to label or stigmatize him.

Another finding worthy of discussion in reviewing the Sample IV Concept of Self data is the consistently higher scores earned by District R pupils as compared with District A pupils of comparable ability. In every one of the 15 comparisons made between pupils of comparable ability over the three administrations of the IAV, District R pupils obtained the mean higher score and in 11 of these comparisons, this difference was statistically significant. The fact that these differences were slight at the fifth grade level and fairly large at the seventh grade level suggests that they cannot be attributed to initial differences in the pupils themselves but are more likely due to the grouping treatments. Treatment effects seem by far the most likely explanation as it will be recalled that the pupils in each district came from several different schools and were found to be reasonably comparable on variables such as socioeconomic status that could have had an effect upon Concept of Self scores.

### Acceptance of Self Data

Bills considers the Acceptance of Self score to be of major importance in interpreting the IAV. It should be noted, however, that correlations between Concept of Self (Col. 1), Acceptance of Self (Col. 2) and the Discrepancy Score (Col. 4) on the adult form of the IAV are quite high. Bills reports correlations in the test manual based upon 300 college students of .90 between Concept of Self and Acceptance of Self; -.83 between Concept of Self and Discrepancy; and -.67 between Acceptance of Self and Discrepancy.[1] It was doubtful that the correlations between these measures would be equally high on the elementary or junior high school forms of the test as these forms have somewhat lower reliabilities than the adult form. In any event, however, the reader is safe in assuming that the three scores measure closely related aspects of the individual's total self-concept.

Reference to Table 10-8 reveals patterns on the Acceptance of Self scores that are generally similar to those found on the Concept of Self scores shown in Table 10-7.

It will be noted that District R pupils generally received higher Acceptance of Self scores than District A pupils of comparable ability. Between-treatment comparisons over the three replications revealed that all three of the first year differences were statistically significant beyond the .01 level of confidence in favor of District R pupils. For the last two years, 12 between-treatment comparisons were made. Five of the six differences between boy's subgroups were statistically significant, while none of the differences between girl's subgroups reached significance. It is interesting to speculate on the possible reasons for the greater between-treatment differences between boy's groups. This difference also shows up, but to a lesser degree in Table 10-7.

The main difference found between Acceptance of Self results and Concept of Self results is the lack of significant level differences found in the former. Only one statistically significant within-treatment difference was found, suggesting that ability level had little or no relationship to the individual's self acceptance in either the ability grouped or random grouped classes. When these results are coupled with the rather consistent between-treatment differences favoring District R one must conclude that ability grouping presents the male pupil with a situation in which his self acceptance is likely to be lower regardless of his ability level. It is possible that the lower self acceptance of superior and average District A pupils as compared with their District R counterparts is due to the greater competition and the presence of fewer pupils of inferior ability in the ability grouping situation. Thus, the bright or average pupil may be less likely to be self-accepting concerning his performance than would be the case in District R where these pupils can usually point to others in the class who are achieving less. If this line of reasoning is followed, however, one would expect District A slow pupils to be more self-acceptant as they would be more likely to hold a relatively more favorable position within the ability grouped classroom. In general, the differences between District A and District R were smallest

---

[1] Ibid., p. 62.

Table 10-8

Sample IV - Differences in Acceptance of Self Over a Three Year Period

| | 5th Grade | | Significant Differences | | 6th Grade | | Significant Differences | | 7th Grade | | Significant Differences | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N[a] | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 129 | 45.47 | | | 54 | 85.52 | | | 43 | 88.77 | | |
| 2. Superior Girls | | | | | 57 | 91.26 | | | 40 | 89.60 | | |
| 3. Average Boys | 211 | 47.19 | | | 100 | 87.64 | | | 104 | 88.62 | | |
| 4. Average Girls | | | | | 97 | 89.91 | | | 106 | 90.98 | | |
| 5. Slow Boys | 80 | 47.16 | | | 39 | 86.26 | | | 26 | 86.88 | | |
| 6. Slow Girls | | | | | 18 | 88.89 | | | 15 | 91.13 | | |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 207 | 49.54 | gr.1-2<br>3.84 | | 140 | 90.51 | gr.1[b]<br>2.73 | | 119 | 94.13 | gr.1<br>3.07 | |
| 8. Superior Girls | | | | | 146 | 89.73 | | | 111 | 91.82 | | |
| 9. Average Boys | 189 | 50.13 | gr.3-4<br>3.25 | | 99 | 92.00 | gr.3<br>2.93 | gr.11<br>2.11 | 116 | 95.00 | gr.3<br>4.91 | |
| 10. Average Girls | | | | | 107 | 90.54 | | | 114 | 92.62 | | |
| 11. Slow Boys | 129 | 51.24 | gr.5-6<br>2.97 | | 89 | 88.66 | | | 53 | 92.55 | gr.5<br>2.06 | |
| 12. Slow Girls | | | | | 41 | 91.93 | | | 42 | 94.26 | | |

[a] Groups not divided by sex for 5th grade data.

[b] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

for pupils in the slow groups. Fewer of these differences were statistically significant and were significant at a lower level of confidence than was found in comparisons of superior and average pupils. These differences, however, still favored the District R pupil rather than the District A pupil. It may be that the gain in self-acceptance that the ability grouped slow pupil may make as a result of being placed in a classroom where he is better able to compete is more than counteracted by the loss of self-acceptance he experiences as a result of the stigma attached to membership in the slow group. This rationale was followed in setting up the working hypothesis.

## Discrepancy Score Data

Although the Discrepancy Score, obtained in this study by subtracting the Concept of Self score from the Ideal Self score, is the last score obtained in the IAV sequence, we will treat it at this time because of it's apparent similarity to the Acceptance of Self score. It would appear that the Discrepancy Score measures self acceptance somewhat less directly and in a different manner than that employed by the Acceptance of Self score itself. In the other scores obtained, the higher scores are indicative of a more favorable self-concept. The reverse is true, however, for the Discrepancy Score as a large score indicates a great discrepancy between the way the individual views himself and the ideal which he would like to attain. Thus, in interpreting Table 10-9, the reader should keep in mind that the higher the Discrepancy Score, the less favorably the individual perceives himself as compared with an ideal. It should be recalled in interpreting the Discrepancy Score, however, that although it is aimed at substantially the same aspect of self-concept as the Acceptance of Self score, the latter is based on Column 2 while the Discrepancy Score is based on the difference between Column 1 and Column 3 of the IAV. Thus, in effect, we have two different measures of self-acceptance, which although similar, are sufficiently different so that the results of each provides something unique to the total picture.

Two analyses were carried out on the Discrepancy Score data. Examination of the results of the first analysis showed that in the majority of cases the low ability pupils had smaller Discrepancy Scores which would indicate better adjustment. As these data seemed to indicate that slow pupils had the best adjustment to their self-concept, and as this part of the Discrepancy Score results did not agree with the Acceptance of Self data presented in Table 10-8, it was decided that a careful recheck should be made of the IAV test papers to see if some spurious element had entered the situation to reduce the size of Discrepancy Score for low ability pupils. A number of cases were found upon examination of Sample IV scores that suggested that pupils had mechanically checked the responses on the various columns of the IAV. In other words, pupils had followed identical patterns in checking responses on the three columns which usually suggests either a negative attitude concerning the test or a lack of understanding concerning the instructions. Other papers were found in which pupils received a higher score in the Concept of Self column than in the Ideal Self column, thus yielding a negative Discrepancy Score. Bills recommends calculating the Discrepancy Score by taking the difference between the Column 1 and and Column 3 entry, regardless of which is larger. This procedure was not followed as it seemed likely that persons whose Discrepancy Scores were negative, i.e., Concept of Self higher than Ideal Self would not be comparable to pupils with equal but positive Discrepancy Scores. Most of these papers, yielding negative Discrepancy Scores seemed to indicate a misunderstanding of the instructions, while a few

Table 10-9

Sample IV - Differences in Discrepancy Scores Over a Three Year Period

| Group | 5th Grade Significant Differences | | | | 6th Grade Significant Differences | | | | 7th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N[a] | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 117 | 6.93 | gr.7-8<br>3.79 | | 53 | 17.53 | gr.7[b]<br>5.11 | gr.3,5<br>2.07;3.06 | 42 | 13.98 | gr.7<br>4.00 | |
| 2. Superior Girls | | | | | 55 | 14.05 | | | 40 | 12.15 | gr.8<br>A.S. | |
| 3. Average Boys | 181 | 6.69 | gr.9-10<br>2.43 | | 78 | 14.28 | gr.9<br>2.28 | | 87 | 13.94 | gr.9<br>3.14 | |
| 4. Average Girls | | | | | 78 | 14.05 | | | 93 | 15.41 | gr.10<br>4.08 | gr.2<br>2.23 |
| 5. Slow Boys | 65 | 6.33 | | | 23 | 11.30 | | | 15 | 13.80 | | |
| 6. Slow Girls | | | | | 13 | 12.23 | gr.12<br>2.30 | | 11 | 17.00 | gr.12<br>2.80 | gr.2<br>2.13 |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 167 | 5.29 | | | 128 | 11.25 | | | 109 | 9.39 | | |
| 8. Superior Girls | | | | | 136 | 12.77 | | gr.12<br>3.70 | 104 | 10.06 | | |
| 9. Average Boys | 160 | 5.73 | | | 85 | 11.26 | | | 95 | 10.29 | | |
| 10. Average Girls | | | | | 90 | 12.46 | | gr.12<br>3.17 | 101 | 10.74 | | |
| 11. Slow Boys | 75 | 5.72 | | | 53 | 9.41 | | | 34 | 10.53 | | |
| 12. Slow Girls | | | | | 24 | 5.67 | | | 26 | 8.62 | | |

[a] Groups not divided by sex for 5th grade data.

[b] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

appeared to identify pupils who consciously reported themselves superior to their own ideal self. The apparently spurious papers were more frequent from pupils of low ability and as such papers yielded a zero or negative Discrepancy Score, they would tend to lower disproportionately the mean Discrepancy Score for low ability pupils. In order to remove these apparently spurious papers, a set of criteria were developed for identifying probably spurious papers and these were removed.

The percentage of papers removed by year and ability level are given in Table 10-10. It may be seen that pupils of comparable ability level in the two districts did not differ materially in the percentage of apparently spurious papers with the exception of a rather large difference between superior pupils during the first year of the study. Overall percentages, however, indicate that 8 per cent of the superior pupils, 15 per cent of the average pupils and 30 per cent of the slow pupils had papers that were defined as spurious based on the criteria established. After these papers were removed, a second analysis of variance of the Discrepancy Score data was calculated. The results of this recalculation may be found in Table 10-9. Although the t values differed somewhat, the same between-treatment differences were statistically significant in the same direction in the two analyses with one exception. During year two, data obtained after spurious scores were removed showed a significantly higher Discrepancy Score for average boys in District A. This difference was not statistically significant for the data obtained in the original analysis, although the difference approached significance in the same direction.

A number of changes, however, occurred in the within-treatment comparisons. These differences reflected generally smaller within-treatment differences after the apparently spurious scores had been removed. Even after removing these scores, however, it will be noted that slow pupils generally had lower mean Discrepancy Scores. As the data obtained in the second analysis of Discrepancy Scores appeared more likely to be correct and was more consistent with the results on the other subtest scores, interpretation of the results were based on that analysis.

In reviewing the data presented in Table 10-9, it will be noted that District R pupils received more desirable (i.e., lower) Discrepancy Scores than District A pupils of comparable ability in all 15 comparisons made in the three replications of the study. Ten of these 15 differences are statistically significant. Four of the five differences between superior pupils in the two districts reached statistical significance. It will be noted that the significant differences occur most frequently between superior pupils and least frequently in comparisons between slow pupils. This is probably due in part to the small numbers of cases in the slow groups, as the actual differences between the mean Discrepancy Scores of the slow groups in the two districts is quite large except at the fifth grade level.

An important trend reflected in the between-treatment difference was that these differences became greater over the three year period of the study. Four of the six between-treatment differences during the third year were statistically significant and the t's for each ability level were higher than the t's for the same ability levels during the first year. Thus, a differential effect favoring District R seemed to be operating with respect to changes in the Discrepancy Score. Comparisons of the magnitude of these scores from year to year can best be made between the final two years as the Elementary School IAV administered during the fifth grade is shorter and yields scores that are not closely comparable to those obtained using the Junior High School IAV in the sixth and seventh grades. The above findings in general support those already discussed in regard to the Acceptance of Self scores.

Table 10-10

Sample IV - Zero and Negative Discrepancy Scores by Ability Level and Treatment

| | 5th Grade Percentage | 6th Grade Percentage | 7th Grade Percentage | Overall Percentage |
|---|---|---|---|---|
| **District A** | | | | |
| 1. Superior | 9 | 3 | 1 | 4.98 |
| 2. Average | 14 | 21 | 14 | 16.34 |
| 3. Slow | 19 | 37 | 38 | 28.98 |
| **District R** | | | | |
| 4. Superior | 19 | 8 | 7 | 10.96 |
| 5. Average | 15 | 15 | 15 | 15.06 |
| 6. Slow | 20 | 41 | 37 | 33.75 |

## Ideal Self Data

The Ideal Self score on the Index of Adjustment and Values calls for the individual to construct an ideal based on the adjectives presented in the scale. In other words, the pupil reports the type of person he would like to be; thus reflecting his aspirations in terms of future self. It was anticipated in this project that the Ideal Self measure would be less likely to be influenced by ability grouping or random grouping than the other scores obtained from the IAV. It may be seen in Table 10-11 that in general this expectation was realized in the results. In reviewing the between-treatment differences as shown in Table 10-11, it will be seen that the mean differences between pupils of comparable ability in the two districts were small during all three years in which the IAV was administered. No significant between-treatment differences were found during the first year. Two significant differences, between the two average groups, were found in the subsequent year. It is interesting to note that for both years two and three, District R average boys obtained significantly higher Ideal Self scores than comparable District A pupils.

During the first year, only one within-treatment difference reached statistical significance but these differences were generally large during the subsequent two years. In District A, seven of the twelve comparisons were statistically significant, six beyond the .01 level. In District R, 11 of these 12 differences were statistically significant--10 beyond the .01 level.

From the above data, one may conclude that Ideal Self as measured by the IAV, was influenced very little by the differences that existed in ability grouped classrooms as compared with random grouped classrooms. A possible exception to this conclusion would be the average pupils who tended to have somewhat higher Ideal Self scores in District R. Level differences were consistent, and quite large during the last two years of the study, indicating that the Ideal Self score was positively related to ability level and that this relationship operated in about the same fashion whether pupils were ability grouped or not.

## Self Concept Changes Accompanying Changes in Ability Level in District A

In view of previous research that indicates most pupils are aware of the ability level to which they are assigned in ability grouped systems, it was hypothesized that reassignment of District A pupils to a lower ability section would be accompanied by changes in the Concept of Self, Acceptance of Self, and Discrepancy Scores on the IAV. In order to test this hypothesis at the elementary school level, pupils in Sample IV who had completed the IAV in both the fifth and sixth grades were classified into three groups: (1) pupils who were assigned to the same ability levels in both grades, (2) pupils who were assigned to a higher ability level in the sixth grade than they had occupied in the fifth grade and (3) pupils who were assigned a lower ability level in the sixth grade than they occupied in the fifth grade. It may be seen in Table 10-12 that the vast majority of pupils were assigned to the same level during the fifth and sixth grade. A total of 26 pupils for whom complete data were available moved to a higher ability level and 37 pupils moved to a lower level.

These numbers, although smaller than would be desirable, were considered large enough to justify a comparative analysis of the self-concept data. The fifth grade and sixth grade mean scores on the four IAV subtests earned by these three groups of pupils were compared using analysis of variance and then the sixth grade means were

## Table 10-11

Sample IV - Differences in Ideal Self Scores Over a Three Year Period

| | 5th Grade Significant Differences | | | | 6th Grade Significant Differences | | | | 7th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N[a] | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 129 | 50.23 | | gr.5-6<br>2.24 | 54 | 100.27 | | gr.3,5[b]<br>4.90;7.69 | 43 | 101.86 | | gr.3,5<br>4.11;6.50 |
| 2. Superior Girls | | | | | 57 | 101.16 | | gr.4,6<br>3.40;3.13 | 40 | 103.15 | | |
| 3. Average Boys | 211 | 49.60 | | | 100 | 91.92 | | | 104 | 95.31 | | gr.5<br>2.25 |
| 4. Average Girls | | | | | 97 | 95.30 | | | 106 | 100.83 | | |
| 5. Slow Boys | 80 | 49.20 | | | 39 | 88.40 | | | 26 | 90.19 | | |
| 6. Slow Girls | | | | | 18 | 94.94 | | | 15 | 97.20 | | |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 207 | 49.86 | | | 140 | 100.24 | | gr.9,11<br>2.37;8.52 | 119 | 101.93 | | gr.9,11<br>2.89;6.93 |
| 8. Superior Girls | | | | | 146 | 101.83 | | gr.10,12<br>3.06;7.26 | 111 | 102.59 | | gr.12<br>6.39 |
| 9. Average Boys | 184 | 49.68 | | | 99 | 97.97 | gr.3<br>4.20 | gr.11<br>5.90 | 116 | 99.97 | gr.3<br>4.15 | gr.11<br>4.87 |
| 10. Average Girls | | | | | 107 | 99.24 | gr.4<br>2.91 | gr.12<br>4.74 | 114 | 101.96 | | gr.12<br>5.82 |
| 11. Slow Boys | 94 | 49.65 | | | 89 | 88.75 | | | 53 | 92.77 | | |
| 12. Slow Girls | | | | | 41 | 91.88 | | | 42 | 95.98 | | |

[a] Groups not divided by sex for fifth grade data.

[b] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

Table 10-12

Sample IV - Changes in Ability Level, 5th to 6th Grade in District A Sample, and Accompanying Changes in IAV Scores

| Group | N | Concept of Self |  | Significant Differences | Acceptance of Self |  | Significant Differences | Ideal Self |  | Significant Differences | Discrepancy Score |  | Significant Differences |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | 5th Grade Score | 6th Grade Score |  | 5th Grade Score | 6th Grade Score |  | 5th Grade Score | 6th Grade Score |  | 5th Grade Score | 6th Grade Score |  |
| 1. Pupils assigned same ability level, 5th and 6th grades | 248 | 44.25 | 85.05 |  | 46.77 | 89.03 |  | 49.72 | 95.95 |  | 5.93 | 11.93 |  |
| 2. Pupils assigned higher ability level in 6th grade | 26 | 44.04 | 84.15 |  | 45.73 | 89.04 |  | 50.11 | 95.04 |  | 6.08 | 12.43 |  |
| 3. Pupils assigned lower ability level in 6th grade | 37 | 42.00 | 81.43 |  | 45.86 | 85.70 |  | 49.43 | 96.67 |  | 7.24 | 15.32 |  |

adjusted for fifth grade differences using analysis of covariance to determine whether significant changes in the self-concept scores had taken place.

## Concept of Self

In reviewing the Concept of Self scores, it is found that groups 1 and 2 (pupils assigned to the same or higher level in the sixth grade) obtained significantly higher scores than did group 3 pupils on the fifth grade administration of the IAV. This difference, although statistically significant at the .05 level, was small. At the sixth grade level, group 3 again scored significantly lower than the other two groups. Data obtained from the covariance analysis indicated that group 3 had the lowest adjusted final mean score but this mean was not significantly lower than adjusted means obtained by the other two groups.

These data indicate that pupils assigned to a lower ability level in the sixth grade were significantly lower in Concept of Self both before and after this reassignment, but that the reassignment itself did not appear to lead to any marked deterioration of the individual's Concept of Self scores. Thus, the lower Concept of Self score may have been a contributing cause to the reassignment of the pupil to a lower ability level or may have resulted from initial misassignment of the pupil to a level where he could not compete. Little evidence is present in the Sample IV data, however, to indicate that assignment of the pupil to a lower ability level brings about a lowering of Concept of Self.

## Acceptance of Self

The Acceptance of Self data show fifth grade scores for the three groups to be very similar, with none of the differences between these initial scores reaching statistical significance. Although no significant differences were found between either the sixth grade means or the adjusted sixth grade means, a trend in the hypothesized direction was apparent as groups 1 and 2 obtained almost identical final scores while group 3 was somewhat lower. The difference between these adjusted final sixth grade means approached statistical significance at the .10 level ($t = 1.54$) for the comparison between groups 1 and 3.

## Ideal Self

It was tentatively hypothesized that no changes in Ideal Self scores would take place for pupils assigned to the same ability level in subsequent years but that assignment to a lower ability level would tend to lower Ideal Self scores and assignment to higher ability level would tend to raise Ideal Self scores. This hypothesis was not supported by the data as may be seen in Table 10-12. Both fifth grade and sixth grade group means on the Ideal Self score were very similar and none of the differences obtained approached statistical significance.

## Discrepancy Score

The mean Discrepancy Score, essentially an indirect measure of self acceptance, was expected to go up for group 2 and go down for group 3 on the sixth grade testing. Reference to Table 10-12 shows a trend for group 3 pupils to have higher Discrepancy

Scores on both the fifth and sixth grade administrations of the IAV. The differences at the fifth grade level were very slight, but the sixth grade differences approached statistical significance. The sixth grade mean after adjustment for fifth grade differences using covariance analysis did not reveal statistically significant differences.

In summary, the above evidence suggests that Concept of Self differences may operate as a contributing cause to the individual in District A being assigned to a lower ability level. Actual reassignment to a lower ability level seems to have a small but consistent effect upon Acceptance of Self and Discrepancy Scores. The small number of pupils who were shifted in ability level, however, make the above conclusions highly tentative.

## SAMPLE VI RESULTS

### Introduction

Pupils in Sample VI were administered the junior high school form of the Index of Adjustment and Values in the seventh, eighth, and ninth grades. The IAV results for this sample provide the best evidence available in this study of the effects of ability grouping upon self-concept at the junior high school level.

### Sex Differences

Comparisons between boys and girls in the same grouping treatment and at the same ability level for Sample VI are given in Table 10-13. It will be noted that of the nine comparisons between the sexes on the Concept of Self score for District A, only one difference, between superior boys and girls, reached statistical significance. These data, however, do reflect a tendancy for girls to receive higher Concept of Self scores than boys, particularly during the last two years of the study. This difference is much more pronounced in the sex comparisons for District R. All nine of these comparisons show girls to have higher Concept of Self means than boys with seven of these differences reaching statistical significance.

A comparison between the sexes on the Acceptance of Self score revealed no significant differences in District A and only two, both favoring superior girls over superior boys, in the nine comparisons for District R. In District R, it will be noted that there was a consistent trend for girls to receive higher Acceptance of Self scores than boys. No such trend was apparent for District A data, however.

On the 18 between-sex comparisons on Ideal Self, it will be noted in Table 10-13 that girls had higher mean scores in all cases. In District A, these differences were fairly small with only two of the nine comparisons reaching statistical significance. In District R, however, the sex differences were generally somewhat larger with seven of the nine differences being statistically significant at or beyond the .05 level.

No significant sex differences in the Discrepancy Score were found in either disstict on the eighth or ninth grade data. During the final year of the study, however, there appeared to be some tendancy for boys to obtain higher Discrepancy Scores than girls. Two of the six sex comparisons made on the final year's data were statistically significant with boys obtaining the higher mean scores in both cases.

## Table 10-13

### Sex Differences on Junior High School IAV Scores for Sample VI

| Groups | 1959-60 IAV 7th Grade $M_{Boys}$ | $M_{Girls}$ | t | 1960-61 IAV 8th Grade $M_{Boys}$ | $M_{Girls}$ | t | 1961-62 IAV 9th Grade $M_{Boys}$ | $M_{Girls}$ | t |
|---|---|---|---|---|---|---|---|---|---|
| Concept of Self | | | | | | | | | |
| District A | | | | | | | | | |
| Superior Boys vs. Superior Girls | 84.21 | 86.58 | N.S. | 88.48 | 89.98 | N.S. | 89.39 | 93.06 | 2.82 |
| Average Boys vs. Average Girls | 84.75 | 83.83 | N.S. | 87.74 | 87.94 | N.S. | 88.94 | 89.90 | N.S. |
| Slow Boys vs. Slow Girls | 85.63 | 85.09 | N.S. | 87.22 | 88.55 | N.S. | 88.92 | 90.05 | N.S. |
| District R | | | | | | | | | |
| Superior Boys vs. Superior Girls | 86.86 | 89.72 | N.S. | 86.68 | 91.24 | 3.72 | 90.54 | 93.95 | 2.40 |
| Average Boys vs. Average Girls | 85.48 | 88.94 | 2.20 | 87.94 | 91.45 | 2.59 | 89.89 | 94.07 | 3.17 |
| Slow Boys vs. Slow Girls | 85.09 | 86.63 | N.S. | 85.04 | 91.17 | 2.98 | 87.95 | 91.69 | 2.54 |
| Acceptance of Self | | | | | | | | | |
| District A | | | | | | | | | |
| Superior Boys vs. Superior Girls | 87.94 | 88.38 | N.S. | 90.42 | 88.76 | N.S. | 91.18 | 91.14 | N.S. |
| Average Boys vs. Average Girls | 88.40 | 87.76 | N.S. | 89.59 | 89.52 | N.S. | 91.83 | 91.14 | N.S. |
| Slow Boys vs. Slow Girls | 91.10 | 85.77 | N.S. | 92.19 | 88.36 | N.S. | 91.25 | 91.47 | N.S. |
| District R | | | | | | | | | |
| Superior Boys vs. Superior Girls | 89.55 | 92.31 | N.S. | 87.33 | 91.24 | 2.46 | 87.78 | 92.43 | 2.29 |
| Average Boys vs. Average Girls | 89.90 | 93.36 | A.S.[a] | 89.00 | 91.67 | N.S. | 92.46 | 94.20 | N.S. |
| Slow Boys vs. Slow Girls | 90.13 | 91.39 | N.S. | 90.43 | 94.10 | N.S. | 91.14 | 93.42 | N.S. |
| Ideal Self | | | | | | | | | |
| District A | | | | | | | | | |
| Superior Boys vs. Superior Girls | 100.28 | 102.05 | 2.06 | 100.59 | 101.54 | N.S. | 101.96 | 102.59 | N.S. |
| Average Boys vs. Average Girls | 95.75 | 98.01 | A.S. | 98.69 | 100.25 | A.S. | 100.37 | 101.57 | N.S. |
| Slow Boys vs. Slow Girls | 92.43 | 95.50 | N.S. | 97.46 | 99.36 | N.S. | 99.27 | 102.07 | 2.49 |
| District R | | | | | | | | | |
| Superior Boys vs. Superior Girls | 100.45 | 103.46 | 2.78 | 98.59 | 100.80 | N.S. | 100.74 | 103.23 | 2.41 |
| Average Boys vs. Average Girls | 97.49 | 102.14 | 3.08 | 99.06 | 101.49 | 2.17 | 99.84 | 102.48 | 2.04 |
| Slow Boys vs. Slow Girls | 96.98 | 98.49 | N.S. | 94.71 | 101.47 | 3.62 | 97.90 | 101.42 | 2.79 |
| Discrepancy Score | | | | | | | | | |
| District A | | | | | | | | | |
| Superior Boys vs. Superior Girls | 16.86 | 15.32 | N.S. | 12.04 | 11.48 | N.S. | 12.57 | 9.56 | 2.49 |
| Average Boys vs. Average Girls | 12.54 | 14.22 | N.S. | 11.02 | 12.66 | N.S. | 11.99 | 11.76 | N.S. |
| Slow Boys vs. Slow Girls | 10.10 | 12.04 | N.S. | 12.02 | 10.88 | N.S. | 10.35 | 12.34 | N.S. |
| District R | | | | | | | | | |
| Superior Boys vs. Superior Girls | 13.91 | 13.92 | N.S. | 12.70 | 10.92 | N.S. | 10.58 | 9.29 | N.S. |
| Average Boys vs. Average Girls | 13.63 | 13.30 | N.S. | 12.68 | 10.00 | N.S. | 11.00 | 8.41 | 2.03 |
| Slow Boys vs. Slow Girls | 12.06 | 14.24 | N.S. | 10.48 | 10.30 | N.S. | 10.59 | 9.87 | N.S. |

[a] Approaching Significance - t=1.65 through 1.95.

## Concept of Self Data

Scores for Sample VI pupils on the Concept of Self column of the IAV are summarized in Table 10-14. For the seventh grade administration of the IAV (when Sample VI pupils had been exposed to one year of ability grouping or random grouping at the sixth grade level and most of a year at the junior high school level) two significant between-treatment differences emerged. District R superior girls were found to be significantly higher in Concept of Self than District A superior girls, and District R average girls were significantly higher than girls of comparable ability in District A. At the eighth and ninth grade testing, District R superior girls again made higher scores than District A superior girls, but these differences did not reach statistical significance. In the case of average girls, however, the District R sample was significantly higher than comparable District A girls in all three administrations of the junior high school Index of Adjustment and Values. Of the nine comparisons between girls of comparable ability in Districts A and R, it will be noted that District R girls received higher Concept of Self scores in every comparison. In contrast, there were no significant between-treatment differences for comparable boys' groups in District A versus District R. Of the nine differences, the District R subgroups were higher in four cases and the District A subgroups in the other five. These data suggest that differences in the two grouping treatments had a somewhat greater effect upon girls than upon boys with regard to Concept of Self scores.

Comparisons within each treatment were carried out next in order to find out if ability grouping or random grouping had differential effects upon pupils at the different ability levels. Of the 18 within-treatment comparisons in District A, three differences were statistically significant and one additional difference approached significance. Of these four differences, three were between superior and average girls and reflected higher self-concept among superior girls. The final significant difference was between superior and slow girls and was again in favor of the superior girls. It is interesting to note in these data that District A average girls made lower Concept of Self scores than did District A slow girls in all three replications. Of the 18 comparisons between District R pupils of different ability levels, none were statistically significant although two differences approached significance during the final year of the study.

A review of the within-treatment comparisons on the Concept of Self scores, however, would suggest that the pupil's increased awareness of his or her ability level, which seems inevitable in the ability grouped situation, may lead to the greater differences found in District A. The impact of ability grouping upon self-concept appears to be more strongly felt among girls.

## Acceptance of Self Data

Scores on this section of the IAV again seem to reflect the greater impact of the grouping treatments on girls' subgroups (Table 10-15). The pattern of between-treatment differences is similar to that found in Table 10-14. No significant between-treatment differences were found in the nine comparisons between District R and District A boys of comparable ability. Two of these differences did approach statistical significance; District A superior boys having obtained higher Acceptance of Self scores in both eighth and ninth grades. Differences between comparable girls' groups in the two districts, revealed that of the nine comparisons made, all favored District R pupils. Two of these differences were statistically significant and three more approached significance.

## Table 10-14

Sample VI - Differences in Concept of Self Scores Over a Three Year Period

| | | 7th Grade Significant Differences | | | | 8th Grade Significant Differences | | | | 9th Grade Significant Differences | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 66 | 84.21 | | | 69 | 88.48 | | | 56 | 89.39 | | |
| 2. Superior Girls | 79 | 86.58 | | gr.4[a]<br>2.18 | 88 | 89.98 | | gr.4[b]<br>A.S. | 64 | 93.06 | | gr.4,6<br>2.64;2.14 |
| 3. Average Boys | 157 | 84.75 | | | 131 | 87.74 | | | 105 | 88.94 | | |
| 4. Average Girls | 167 | 83.84 | | | 127 | 87.94 | | | 104 | 89.90 | | |
| 5. Slow Boys | 30 | 85.63 | | | 41 | 87.22 | | | 60 | 88.91 | | |
| 6. Slow Girls | 22 | 85.09 | | | 42 | 88.54 | | | 55 | 90.05 | | |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 58 | 86.86 | | | 111 | 86.68 | | | 95 | 90.54 | | gr.11<br>A.S. |
| 8. Superior Girls | 54 | 89.72 | gr.2<br>2.14 | | 84 | 91.24 | | | 44 | 93.95 | | |
| 9. Average Boys | 83 | 85.48 | | | 66 | 87.56 | | | 46 | 89.89 | | |
| 10. Average Girls | 86 | 88.94 | gr.4<br>4.00 | | 51 | 91.45 | gr.4<br>2.59 | | 60 | 94.07 | gr.4<br>3.49 | gr.12<br>A.S. |
| 11. Slow Boys | 47 | 86.45 | | | 42 | 85.05 | | | 59 | 87.94 | | |
| 12. Slow Girls | 33 | 86.64 | | | 40 | 91.17 | | | 64 | 91.69 | | |

[a] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

[b] A.S. indicates that the t-test is approaching significance - t=1.65 through 1.95.

Table 10-15

Sample VI - Differences in Acceptance of Self Scores Over a Three Year Period

| | 7th Grade Significant Differences | | | | 8th Grade Significant Differences | | | | 9th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 66 | 87.94 | | | 69 | 90.42 | gr.7[b]<br>A.S. | | 56 | 91.18 | gr.7<br>A.S. | |
| 2. Superior Girls | 79 | 88.37 | | | 88 | 88.76 | | | 64 | 91.14 | | |
| 3. Average Boys | 157 | 88.40 | | | 131 | 89.59 | | | 105 | 91.83 | | |
| 4. Average Girls | 167 | 87.76 | | | 127 | 89.52 | | | 104 | 91.25 | | |
| 5. Slow Boys | 30 | 91.10 | | | 41 | 92.19 | | | 60 | 91.25 | | |
| 6. Slow Girls | 22 | 85.77 | | | 42 | 88.35 | | | 55 | 91.47 | | |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 58 | 89.55 | | | 111 | 87.33 | | | 95 | 87.78 | | |
| 8. Superior Girls | 54 | 92.31 | gr.2<br>A.S. | | 84 | 91.24 | | | 44 | 92.43 | | |
| 9. Average Boys | 83 | 89.90 | | | 66 | 89.00 | | | 46 | 92.46 | | gr.7<br>2.36 |
| 10. Average Girls | 86 | 93.36 | gr.4[a]<br>3.39 | | 51 | 91.67 | | | 60 | 94.20 | gr.4<br>A.S. | |
| 11. Slow Boys | 47 | 90.13 | | | 42 | 90.43 | | | 59 | 91.13 | | gr.7<br>A.S. |
| 12. Slow Girls | 33 | 91.39 | gr.6<br>A.S. | | 40 | 94.10 | gr.6<br>2.04 | | 64 | 93.42 | | |

[a] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

[b] Indicates that the t-test is approaching significance - t=1.65 through 1.95.

## Discrepancy Score Data

It will be recalled that Discrepancy Scores in Sample IV were reanalyzed after removing scores that, upon examination of the answer sheets, appeared to be spurious. Such answer sheets were removed from the Sample VI data prior to start of the analysis so that scores for these cases were not included in any of the IAV subtests for Sample VI.

For Sample IV, rather large and consistent Discrepancy Score differences favoring District R were found in the between-treatment analysis. The between-treatment differences in Sample VI generally followed the same pattern but were much smaller with only two differences, both between average girls in the two districts, being statistically significant (see Table 10-16). Differences in Discrepancy Scores for superior boys in the two districts approached significance in the second and fourth years of the study with District R pupils receiving more favorable scores in both instances. It will be noted in Table 10-16 that between-treatment differences reaching or approaching significance increased during the final year of the study. Even in this final year, however, these differences were considerably less marked than were found in Sample IV, suggesting that the grouping treatment at the junior high school level had somewhat less of an effect upon Discrepancy Scores than the more extensive treatment differences that were present at the elementary school level.

Within-treatment differences between Discrepancy Score means show no consistent pattern. In District A both superior boys and girls had somewhat higher Discrepancy Scores than pupils at the other ability levels. However, this pattern was not found for the second and third years' data. In fact, during the third year superior girls had the lowest Discrepancy Score and were significantly lower than both average and slow girls for that year.

Within-treatment comparisons for District R revealed no significant differences among District R pupils on any of the three replications.

## Ideal Self Data

Results for Sample VI on the IAV Ideal Self score are summarized in Table 10-17. Of the nine between-treatment comparisons for boys, six favored District A and three favored District R groups. These differences were small with only one, favoring District R slow boys, being statistically significant. A somewhat more definite pattern emerged, however, in the between-treatment differences for girls. In seven of the nine comparisons, the District R girls scored higher, with two of these differences reaching statistical significance.

As was the case with Sample IV, (see Table 10-11) within-treatment differences on Ideal Self were fairly large and generally favored pupils of higher ability. District A superior boys had significantly higher Ideal Self scores than slow boys for all three replications. District A superior girls received higher Ideal Self scores than average and slow girls in all cases, these differences reaching statistical significance in three of the six comparisons. A similar pattern emerged in the within-treatment comparisons for District R boys with superior boys having a significantly higher mean Ideal Self score than slow boys in all three replications. Superior girls were significantly higher than slow girls in two replications and were also higher than average District R girls during the seventh grade year. During this year, District R average girls were also significantly higher than District R slow girls and in the

Table 10-16

Sample VI - Differences in Discrepancy Scores Over a Three Year Period

| | 7th Grade Significant Differences | | | | 8th Grade Significant Differences | | | | 9th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| District A | | | | | | | | | | | | |
| 1. Superior Boys | 66 | 16.86 | gr.7<br>A.S.[a] | gr.3,5[b]<br>3.02;3.30 | 69 | 12.04 | | | 56 | 12.57 | gr.7<br>A.S. | |
| 2. Superior Girls | 79 | 15.32 | | gr.4<br>A.S. | 88 | 11.48 | | | 64 | 9.56 | | |
| 3. Average Boys | 157 | 12.54 | | | 131 | 11.02 | | | 105 | 11.99 | | |
| 4. Average Girls | 167 | 14.22 | | | 127 | 12.66 | gr.10<br>2.09 | | 104 | 11.76 | gr.10<br>3.08 | gr.2<br>1.96 |
| 5. Slow Boys | 30 | 10.10 | | | 41 | 12.02 | | | 60 | 10.35 | | |
| 6. Slow Girls | 22 | 12.05 | | | 42 | 10.88 | | | 55 | 12.34 | gr.12<br>A.S. | gr.2<br>1.98 |
| District R | | | | | | | | | | | | |
| 7. Superior Boys | 58 | 13.91 | | | 111 | 12.70 | | | 95 | 10.57 | | |
| 8. Superior Girls | 54 | 13.93 | | | 84 | 10.91 | | | 44 | 9.29 | | |
| 9. Average Boys | 83 | 13.63 | | | 66 | 12.68 | | | 46 | 11.00 | | |
| 10. Average Girls | 86 | 13.30 | | | 51 | 10.00 | | | 60 | 8.42 | | |
| 11. Slow Boys | 57 | 12.06 | | | 42 | 10.47 | | | 59 | 10.59 | | |
| 12. Slow Girls | 33 | 14.24 | | | 40 | 10.30 | | | 64 | 9.87 | | |

[a] A.S. indicates that the t-test is approaching significance - t=1.65 through 1.95.

[b] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

### Table 10-17

### Sample VI - Differences in Ideal Self Scores Over a Three Year Period

| Group | 7th Grade Significant Differences | | | | 8th Grade Significant Differences | | | | 9th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | | | | | |
| 1. Superior Boys | 66 | 100.29 | | gr.3,5<br>2.84;5.02 | 69 | 100.59 | | gr.5,3<br>2.18;A.S. | 56 | 101.96 | | gr.5<br>2.41 |
| 2. Superior Girls | 79 | 102.05 | | gr.4,6<br>4.13;4.29 | 88 | 101.54 | | gr.6<br>2.02 | 64 | 102.59 | | gr.4<br>A.S. |
| 3. Average Boys | 157 | 95.75 | | | 131 | 98.69 | | | 105 | 100.37 | | |
| 4. Average Girls | 167 | 98.01 | | | 127 | 100.25 | | | 104 | 101.57 | | |
| 5. Slow Boys | 30 | 92.43 | | | 41 | 97.46 | | | 60 | 99.27 | | |
| 6. Slow Girls | 22 | 95.50 | | | 42 | 99.36 | | | 55 | 102.07 | | |
| **District R** | | | | | | | | | | | | |
| 7. Superior Boys | 58 | 100.45 | | gr.11<br>2.30 | 111 | 98.59 | | gr.11<br>2.21 | 95 | 100.74 | | gr.11<br>2.29 |
| 8. Superior Girls | 54 | 103.46 | gr.2[a]<br>2.05 | gr.10,12<br>A.S.[b];3.98 | 84 | 100.80 | | | 44 | 103.23 | | gr.12<br>2.23 |
| 9. Average Boys | 83 | 97.49 | | | 66 | 99.06 | | gr.11<br>2.55 | 46 | 99.85 | | |
| 10. Average Girls | 86 | 102.14 | gr.4<br>4.32 | gr.12<br>2.92 | 51 | 101.49 | | | 60 | 102.48 | | |
| 11. Slow Boys | 47 | 96.98 | gr.5<br>2.29 | | 42 | 94.71 | | | 59 | 97.90 | | |
| 12. Slow Girls | 33 | 98.48 | | | 40 | 101.47 | | | 64 | 101.42 | | |

[a] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

[b] A.S. indicates that the t-test is approaching significance - t=1.65 through 1.95.

eighth grade, District R average boys were significantly higher than District R slow boys. These within-treatment differences for both districts reflect a fairly strong tendancy for superior pupils to have higher Ideal Self scores than pupils at the average or slow levels.

## Self-Concept Changes Accompanying Ability Level Changes in District A

### Introduction

In order to determine whether the four self-concept variables are affected by changes in the pupil's ability level in ability grouped classes, Sample VI pupils were classified into six groups: boys and girls assigned to the same ability level in the sixth and seventh grades, boys and girls assigned to a higher ability level in the seventh grade, and boys and girls assigned a lower ability level in the seventh grade. Analysis of covariance was carried out for each of the four self-concept variables to determine whether the final adjusted mean scores on these variables were significantly different for like-sex pupils whose ability classification remained the same, went up, or went down. The results of this analysis may be found in Table 10-18. As was the case with Sample IV, the majority of pupils were assigned to the same ability levels in the sixth and seventh grades. A sufficient number of both boys and girls, however, were assigned either higher or lower ability levels to permit an analysis of these data.

### Concept of Self

In comparing the Concept of Self scores for the sixth grade administration of the IAV, it was found that boys assigned the same level (group 1) were significantly higher than boys who were subsequently assigned to a higher level or a lower level. On the seventh grade administration, however, these initial differences had largely disappeared.

In contrast, the three girls' groups were closely comparable on the sixth grade Concept of Self score but differences had emerged by the seventh grade administration. Comparison of the adjusted seventh grade means of girls' groups on the Concept of Self score revealed that girls assigned the same ability level in both sixth and seventh grades (group 2) had higher mean scores than both the girls assigned to a higher level (group 4) and the girls assigned to a lower level in the seventh grade (group 6). This latter difference was significant at the .01 level, and would suggest that in an ability grouped situation, girls find remaining in the same ability level less threatening to self-concept than either moving to a higher level or a lower level, with movement to a lower level having the least favorable effect upon their scores.

### Acceptance of Self

Scores on Acceptance of Self showed that group 1 boys again had a somewhat higher mean score than either group 3 or group 5 boys, but these differences were not statistically significant. Group 2 and group 4 girls scored substantially the same while group 6 girls were significantly lower as compared to group 2 girls. At the seventh grade testing, we find that the three boys' groups had become more

Table 10-18

Sample VI - Changes in Ability Level, 6th to 7th Grade in the District A Sample, and Accompanying Changes in IAV Scores

| Group | N | Concept of Self | | | Acceptance of Self | | | Ideal Self | | | Discrepancy Score | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean Score (6th Grade) | Mean Score (7th Grade) | Significant Differences[a] | Mean Score (6th Grade) | Mean Score (7th Grade) | Significant Differences | Mean Score (6th Grade) | Mean Score (7th Grade) | Significant Differences | Mean Score (6th Grade) | Mean Score (7th Grade) | Significant Differences |
| 1. Boys assigned same ability level, 6th and 7th grades | 154 | 89.90 | 88.04 | | 93.94 | 89.86 | | 103.31 | 99.57 | gr.3<br>3.11 | 14.75 | 11.92 | |
| 2. Girls assigned same ability level, 6th and 7th grades | 157 | 88.19 | 89.33 | gr.6[b]<br>2.63 | 88.75 | 90.10 | gr.6<br>3.40 | 99.50 | 101.20 | gr.6<br>2.77 | 14.30 | 12.05 | |
| 3. Boys assigned higher level in 7th grade | 40 | 83.77 | 85.60 | | 88.07 | 88.45 | | 94.80 | 95.35 | | 13.02 | 10.60 | |
| 4. Girls assigned higher level in 7th grade | 50 | 84.08 | 87.16 | | 86.68 | 87.88 | gr.6<br>2.01 | 97.12 | 99.50 | | 13.18 | 12.60 | |
| 5. Boys assigned lower level in 7th grade | 31 | 85.41 | 87.55 | | 90.81 | 89.74 | | 95.74 | 97.87 | | 11.64 | 10.32 | |
| 6. Girls assigned lower level in 7th grade | 31 | 84.00 | 85.35 | | 83.90 | 82.52 | | 99.32 | 97.32 | | 14.87 | 12.42 | |

---

[a] Significant differences in final means adjusted for differences in initial means, using covariance analysis.

[b] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

closely comparable on this score. A comparison of the adjusted seventh grade means for the girls' groups showed both group 2 and group 4 to have significantly higher adjusted Acceptance of Self means than group 6. Thus, it appears that for girls, being assigned to a higher ability level did not result in a significant increase in Acceptance of Self scores, but assignment to a lower level brought about a relative decline in these scores. These data generally support the Sample IV findings.

### Ideal Self

Initial Ideal Self scores showed the boys subsequently assigned to the same level to be significantly higher than those who were to be assigned to a higher or lower ability level during the subsequent year. Again, the seventh grade data showed a slight drop in Ideal Self for group 1 boys similar to that found in Acceptance of Self data for this group. On the other hand, both group 3 and group 5 showed an increase in Ideal Self scores for the seventh grade means, thus resulting in final adjusted means that were somewhat more comparable than the initial sixth grade means for the boys groups. It will be noted, however, that the group 1 boys were significantly higher than group 3 on the final adjusted Ideal Self mean. Group 3 and group 5 boys were not significantly different on the final adjusted mean.

Comparisons between the three girls' groups showed them to be quite similar in their sixth grade mean scores with somewhat larger differences occurring in the seventh grade means, thus yielding somewhat similar results to those found in the Concept of Self and Acceptance of Self data. The final adjusted mean score for group 2 was significantly higher than that for group 6 on the Ideal Self score. It is interesting to note that although the final adjusted means for group 4 and group 6 girls were not significantly different, the girls who were assigned a higher ability level experienced an increase of about 2 points between the sixth and seventh grade tests while the girls who were assigned a lower level in the seventh grade experienced a decrease of two points on Ideal Self. These data again seem to suggest that for the girls' groups, assignment to a lower ability level tends to lower the individual's Ideal Self score.

### Discrepancy Score

It may be seen that no significant differences in Discrepancy Scores were found among pupils who moved up, down, or remained the same in ability level classification between the sixth and seventh grades. All groups had slightly lower Discrepancy Scores at the seventh grade level.

## SAMPLE VII RESULTS

### Sex Differences

The reader will recall that during the first administration of the elementary school IAV to Sample IV, an analysis was made to determine whether sex differences existed between boys and girls of the same ability level on the four IAV scores. No sex differences emerged from this comparison and as a result, the analysis of the elementary school IAV data combined boys and girls. In order to make a similar check on the junior high school IAV, the first administration of this measure to

Sample VII pupils was analyzed to determine whether sex differences existed between boys and girls of comparable ability level. This analysis revealed no statistically significant sex differences in Concept of Self or Acceptance of Self for the eighth grade administration of the junior high school IAV to Sample VII. (see Table 10-19). Comparisons between boys and girls on Ideal Self, however, showed that girls scored consistently higher on this measure than boys. Three of the six comparisons between the sexes on the Ideal Self score were statistically significant. In both districts superior girls scored significantly higher on Ideal Self than superior boys. In District R, the difference between average boys and girls was also statistically significant. The difference in District R between slow boys and girls approached statistical significance with a t of 1.95.

Comparison of Discrepancy Scores of boys and girls also showed a tendency for girls to be somewhat higher. Two of these Discrepancy Score differences reached statistical significance; those comparing the average boys and girls in each district.

Because the analysis of sex differences in Sample IV, previously discussed, yielded no significantly significant sex differences and, in fact, had yielded very similar mean scores between boys and girls on all of the self-concept variables, it was decided to further check sex differences for Sample VII using the ninth grade data on the junior high school Index of Adjustment and Values. As may be seen in Table 10-19, the results of this analysis were in general agreement with the data obtained during the 1959-60 administration. One significant sex difference was found that favored superior girls in District R on the Concept of Self score. It will be noted that for the ninth grade data a fairly consistent trend for girls to be higher on Concept of Self was present in District R for all ability levels. No such trend existed in District A where, in fact, the superior boys and girls were almost identical and boys were slightly higher than the girls in the average and slow groups. As the differences making up this trend, however, are rather small and the trend is not apparent in the eighth grade data, it appears unlikely that this trend has any real significance. No significant differences were found between boys and girls on Acceptance of Self scores in the ninth grade testing, thus supporting the results obtained in the previous year.

Differences between boys and girls on the Ideal Self score, which were fairly large in the 1959-60 testing were in general found to be even larger in the 1960-61 testing. It will also be noted that the results of the two years are highly consistent. In both years, District A superior girls were higher than boys. In District R, girls were higher on the Ideal Self score at all three ability levels. During the second year, these District R differences were generally larger and all three reached statistical significance.

Sex differences in Discrepancy Scores on the ninth grade testing were found in the average and slow samples in District A, both differences just reaching the .05 level of significance and both favoring girls. These differences are moderately consistent with the eighth grade Discrepancy Score results.

In summary, of the 48 sex difference comparisons shown in Table 10-19, 12 were statistically significant. Five of these significant differences were found in District A and seven in District R. It will be noted that most of these differences were found in the Ideal Self and Discrepancy Score areas. Of the seven significant sex differences found in the Ideal Self area, all showed more favorable scores for girls. There were four significant differences in the between-sex comparisons on the Discrepancy Scores, and in all four cases girls' scores were significantly less

Table 10-19

Sex Differences in Junior High School IAV Scores for Sample VII

| Groups | 8th Grade IAV $M_{Boys}$ | 8th Grade IAV $M_{Girls}$ | 8th Grade IAV t | 9th Grade IAV $M_{Boys}$ | 9th Grade IAV $M_{Girls}$ | 9th Grade IAV t |
|---|---|---|---|---|---|---|
| Concept of Self | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 85.00 | 85.40 | N.S. | 87.70 | 87.74 | N.S. |
| District A - Average Boys vs. District A - Average Girls | 82.67 | 81.47 | N.S. | 85.82 | 84.58 | N.S. |
| District A - Slow Boys vs. District A - Slow Girls | 81.28 | 80.23 | N.S. | 82.92 | 80.50 | N.S. |
| District R - Superior Boys vs. District R - Superior Girls | 86.68 | 88.97 | N.S. | 85.97 | 89.68 | 2.65 |
| District R - Average Boys vs. District R - Average Girls | 83.81 | 83.17 | N.S. | 85.06 | 87.34 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 83.21 | 85.33 | N.S. | 82.32 | 86.15 | N.S. |
| Acceptance of Self | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 88.02 | 84.83 | N.S. | 85.69 | 82.44 | N.S. |
| District A - Average Boys vs. District A - Average Girls | 87.38 | 84.82 | N.S. | 86.40 | 83.53 | N.S. |
| District A - Slow Boys vs. District A - Slow Girls | 84.00 | 83.76 | N.S. | 84.97 | 82.75 | N.S. |
| District R - Superior Boys vs. District R - Superior Girls | 90.62 | 90.15 | N.S. | 87.16 | 87.70 | N.S. |
| District R - Average Boys vs. District R - Average Girls | 86.91 | 87.51 | N.S. | 85.35 | 88.07 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 86.37 | 89.18 | N.S. | 86.82 | 85.84 | N.S. |
| Ideal Self | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 99.79 | 101.47 | 2.21 | 99.64 | 101.69 | 2.47 |
| District A - Average Boys vs. District A - Average Girls | 95.87 | 97.34 | N.S. | 97.19 | 97.92 | N.S. |
| District A - Slow Boys vs. District A - Slow Girls | 87.64 | 88.53 | N.S. | 91.48 | 95.37 | N.S. |
| District R - Superior Boys vs. District R - Superior Girls | 100.45 | 102.44 | 2.54 | 98.75 | 101.92 | 2.99 |
| District R - Average Boys vs. District R - Average Girls | 93.51 | 99.42 | 4.27 | 96.13 | 101.22 | 3.01 |
| District R - Slow Boys vs. District R - Slow Girls | 89.89 | 96.94 | N.S. | 91.82 | 98.69 | 2.25 |
| Discrepancy Score | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 14.96 | 16.22 | N.S. | 11.91 | 13.89 | N.S. |
| District A - Average Boys vs. District A - Average Girls | 13.47 | 17.02 | 2.93 | 12.05 | 14.25 | 2.05 |
| District A - Slow Boys vs. District A - Slow Girls | 9.89 | 11.70 | N.S. | 10.02 | 16.25 | 2.01 |
| District R - Superior Boys vs. District R - Superior Girls | 14.05 | 13.64 | N.S. | 12.76 | 12.25 | N.S. |
| District R - Average Boys vs. District R - Average Girls | 12.33 | 16.27 | 2.22 | 11.17 | 13.73 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 10.29 | 14.00 | N.S. | 9.50 | 12.61 | N.S. |

favorably than boys'. As the Discrepancy Score is developed partially from the Ideal Self score, it appears that all of these differences are traceable to the tendency of girls to obtain somewhat higher Ideal Self scores than boys.

Because of the sex differences found in the data for Sample VII, it was decided to separate the boys and girls for the remainder of the analysis of self-concept data. As the Ideal Self data were generally considered least likely to be influenced by the grouping treatments, it is possible that the decision to divide sexes for this analysis with attendent reduction in the size of the subgroups, does not provide the best interpretation of the data. The analysis is, of course, complicated by dividing the sexes, and the likelihood of making Type II errors is increased because of the aforementioned reduction in subgroup size. It will be noted, however, that the proportion of boys and girls at the different ability levels is far from equal with considerably more girls being classified superior and a greater number of boys being classified slow. This difference is particularly marked in District A, and, does not show up to a comparable degree in District R. Thus, comparisons between groups without regard to sex could lead one to attribute differences to the treatments when in fact, they were caused by differences in proportion of boys and girls in the groups being compared. Thus, it seemed safer to divide each group by sex and this procedure was followed in analysis of the junior high school IAV results for all samples.

Analysis of differences between treatments and within-treatments was carried out with 12 subgroups, separating boys and girls in each ability group, instead of the six groups as was the case with the elementary school IAV data collected on Sample IV pupils.

## Concept of Self Data

In reviewing the between-treatment differences on Concept of Self for Sample VII, we find District R superior girls have significantly higher Concept of Self scores than District A superior girls for the eighth grade testing (see Table 10-20). Other between-treatment differences were not statistically significant although three differences approached statistical significance, all favoring District R girls. Slow District R girls were somewhat higher than girls of comparable ability in District A, this difference approaching statistical significance at the .05 level for both replications. When one considers that only one between-treatment difference reached statistical significance at the .05 level out of the 12 comparisons made, it seems possible that this single significant difference is due to chance. The fact, however, that the three differences which approached significance also favored District R, suggests that if any self-concept difference did exist, it favored the random grouped sample. These between-treatment differences are somewhat similar to those found for Sample VI, in which the four significant differences found all favored District R girls over District A girls. All significant between-treatment differences for Sample IV on the Concept of Self score also favored District R.

With regard to within-treatment differences, five of the twelve comparisons between District A pupils reached statistical significance and two others approached statistical significance. In every case these differences were in favor of the higher ability group. For District R, only two of the twelve within-treatment differences were statistically significant and three others approaching statistical significance at the .05 level. Again, the significant differences favored the group of higher ability.

## Table 10-20

Sample VII - Differences in Concept of Self Scores Over a Two-Year Period

| | 8th Grade | | | | 9th Grade | | | |
|---|---|---|---|---|---|---|---|---|
| | | Significant Differences | | | | Significant Differences | | |
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | |
| 1. Superior Boys | 48 | 85.00 | | gr.5<br>A.S.[a] | 47 | 87.70 | | gr.5<br>2.37 |
| 2. Superior Girls | 59 | 85.41 | | gr.4,6<br>2.44;2.17 | 63 | 87.75 | | gr.4,6<br>2.02;3.04 |
| 3. Average Boys | 135 | 82.67 | | | 128 | 85.82 | | gr.5<br>A.S. |
| 4. Average Girls | 116 | 81.47 | | | 107 | 84.58 | | |
| 5. Slow Boys | 36 | 81.28 | | | 39 | 82.92 | | |
| 6. Slow Girls | 17 | 80.23 | | | 16 | 80.50 | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 79 | 86.68 | | gr.11,9<br>1.96;A.S. | 69 | 85.97 | | gr.11<br>A.S. |
| 8. Superior Girls | 59 | 88.97 | gr.2[b]<br>2.28 | gr.10,12<br>3.77;A.S. | 56 | 89.68 | | |
| 9. Average Boys | 57 | 83.81 | | | 51 | 85.06 | | |
| 10. Average Girls | 59 | 83.17 | | | 53 | 87.34 | gr.4<br>A.S. | |
| 11. Slow Boys | 38 | 83.21 | | | 28 | 82.32 | | |
| 12. Slow Girls | 33 | 85.33 | gr.6<br>A.S. | | 26 | 86.15 | gr.6<br>A.S. | |

[a] A.S. indicates that the t-test is approaching significance - t=1.65 through 1.95.

[b] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

Table 10-21

Sample VII - Difference in Acceptance of Self Scores Over a Two-Year Period

| | | 8th Grade Significant Differences | | | | 9th Grade Significant Differences | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| District A | | | | | | | | |
| 1. Superior Boys | 48 | 88.02 | | | 47 | 85.68 | | |
| 2. Superior Girls | 59 | 84.83 | | | 63 | 82.44 | | |
| 3. Average Boys | 135 | 87.38 | | | 128 | 86.40 | | |
| 4. Average Girls | 116 | 84.82 | | | 107 | 83.53 | | |
| 5. Slow Boys | 36 | 84.00 | | | 39 | 84.94 | | |
| 6. Slow Girls | 17 | 83.76 | | | 16 | 82.75 | | |
| District R | | | | | | | | |
| 7. Superior Boys | 79 | 90.62 | | gr.9<br>A.S.[a] | 69 | 87.16 | | |
| 8. Superior Girls | 59 | 90.15 | gr.2[b]<br>2.51 | | 56 | 87.69 | gr.2<br>2.34 | |
| 9. Average Boys | 57 | 86.91 | | | 51 | 85.35 | | |
| 10. Average Girls | 59 | 87.50 | | | 53 | 88.07 | gr.4<br>2.16 | |
| 11. Slow Boys | 38 | 86.37 | | | 28 | 86.82 | | |
| 12. Slow Girls | 33 | 89.18 | gr.6<br>A.S. | | 26 | 85.84 | | |

[a] A.S. indicates that the t-test is approaching significance - t=1.65 through 1.95.

[b] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

Table 10-22

Sample VII - Differences in Discrepancy Scores Over a Two-Year Period

| Group | 8th Grade N | 8th Grade Mean Score (JHS - IAV) | 8th Grade Significant Differences: Between Treatment | 8th Grade Significant Differences: Within Treatment | 9th Grade N | 9th Grade Mean Score (JHS - IAV) | 9th Grade Significant Differences: Between Treatment | 9th Grade Significant Differences: Within Treatment |
|---|---|---|---|---|---|---|---|---|
| **District A** | | | | | | | | |
| 1. Superior Boys | 48 | 14.96 | | gr.5[b]<br>2.51 | 47 | 11.91 | | |
| 2. Superior Girls | 59 | 16.22 | gr.8[a]<br>A.S. | | 63 | 13.89 | | |
| 3. Average Boys | 135 | 13.47 | | gr.5<br>A.S. | 128 | 12.05 | | |
| 4. Average Girls | 116 | 17.03 | | | 107 | 14.24 | | |
| 5. Slow Boys | 36 | 9.89 | | | 39 | 10.02 | | |
| 6. Slow Girls | 17 | 11.71 | | | 16 | 16.25 | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 79 | 14.05 | | gr.11<br>2.08 | 69 | 12.77 | | gr.11<br>A.S. |
| 8. Superior Girls | 59 | 13.64 | | | 56 | 12.25 | | |
| 9. Average Boys | 57 | 12.33 | | | 51 | 11.18 | | |
| 10. Average Girls | 59 | 16.27 | | | 53 | 13.74 | | |
| 11. Slow Boys | 38 | 10.29 | | | 28 | 9.50 | | |
| 12. Slow Girls | 33 | 14.00 | | | 26 | 12.61 | | |

[a] A.S. indicates that the t-test is approaching significance - t=1.65 through 1.95.

[b] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

Table 10-23

Sample VII - Differences in Ideal Self Scores Over a Two-Year Period

| | 8th Grade Significant Differences | | | | 9th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | |
| 1. Superior Boys | 48 | 99.80 | | gr.3,5<br>3.09;6.36 | 47 | 99.64 | | gr.5,3 [b]<br>4.42;A.S. |
| 2. Superior Girls | 59 | 101.47 | | gr.2,6<br>2.67;6.24 | 63 | 101.70 | | gr.4,6<br>2.37;4.16 |
| 3. Average Boys | 135 | 95.87 | | gr.5<br>4.69 | 128 | 97.19 | | gr.5<br>3.43 |
| 4. Average Girls | 116 | 97.34 | | gr.6<br>2.83 | 107 | 97.92 | | |
| 5. Slow Boys | 36 | 87.64 | | | 39 | 91.49 | | |
| 6. Slow Girls | 17 | 88.53 | | | 16 | 95.37 | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 79 | 100.45 | | gr.9,11<br>3.82;4.59 | 69 | 98.75 | | gr.11<br>3.36 |
| 8. Superior Girls | 59 | 102.44 | | gr.10,12<br>3.59;4.27 | 56 | 101.93 | | gr.12<br>2.33 |
| 9. Average Boys | 57 | 93.51 | | | 51 | 96.14 | | |
| 10. Average Girls | 59 | 99.42 | | | 53 | 101.23 | gr.4<br>A.S. | |
| 11. Slow Boys | 38 | 89.89 | | | 28 | 91.82 | | |
| 12. Slow Girls | 33 | 96.94 | gr.6[a]<br>2.50 | | 26 | 98.69 | | |

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

## Acceptance of Self Data

Between-treatment differences on the Acceptance of Self score for Sample VII revealed three significant differences, all favoring District R girls as compared with District A girls (see Table 10-21). For both administrations of the IAV, District R superior girls were significantly higher than District A superior girls. During the 1960-61 year, District R average girls were significantly higher than District A average girls. The relatively small number of slow girls in the two samples led to a considerably larger standard error of the difference between means for these groups, thus requiring greater differences to reach statistical significance. A rather steady trend appeared to be present in all samples for District R slow girls to score higher on Acceptance of Self than District A slow girls. These results suggest that a true difference is present although this difference reaches statistical significance in only about half of the replications.

With regard to within-treatment differences in self acceptance, none of the differences within District A or District R were statistically significant. These results agree substantially with the findings of other samples in the study.

## Discrepancy Score Data

Comparisons between the mean Discrepancy Scores of subgroups in the two districts revealed no statistically significant between-treatment differences.

Only two statistically significant differences within the treatments were found. In both districts, during the eighth grade year, superior boys made significantly higher Discrepancy Scores than slow boys. This difference also approached significance in District R during the ninth grade year. It will be recalled that apparently spurious answer sheets had been removed from the Sample VII data before analysis. These results suggest that a true difference existed between these groups.

## Ideal Self Data

Between-treatment comparisons for Sample VII on the Ideal Self score revealed only one statistically significant difference favoring District R slow girls.

Comparisons showed that for District A, 10 of the 12 within-treatment differences were statistically significant, all but one beyond the .01 level. These differences in every case favored the pupils of higher ability level. Of the 12 within-treatment comparisons made for District R on Ideal Self scores, six were statistically significant with five of these significant beyond the .01 level. Again, these differences in every case favored the group of higher ability. The greatest differences in Ideal Self scores for both districts were found between the superior and slow pupils (see Table 10-23). The high Discrepancy Score for superior pupils was probably due in great part to their higher Ideal Self scores. These data would suggest that Ideal Self is related to ability but that this relationship is not altered materially by the type of grouping. The presence of only one significant between-treatment difference tends to further support this conclusion.

## SAMPLE VIII RESULTS

### Sex Differences

Because of sex differences found in the IAV scores of Samples VI and VII, Sample VIII pupils were divided by sex and an analysis of variance aimed at comparing differences between treatments, within treatments, and between sexes was carried out following the same procedures used with other samples.

Sample VIII sex differences may be found in Table 10-24. It will be noted that for the Concept of Self score only one statistically significant difference was found. There was also only one significant difference between the sexes on Acceptance of Self. For the Sample VIII data, there were six statistically significant sex differences on the Ideal Self score, in every case reflecting higher Ideal Self scores for girls. Three of the sex differences on the Discrepancy Score were significant, and again, as was the case with Sample VII, these reflected higher (less favorable) Discrepancy Scores for girls. These data, closely following as they do the pattern established in Sample VII, tend to provide fairly reliable evidence that girls at the secondary level are consistently higher in their Ideal Self aspirations and this difference is also reflected in the Discrepancy Score, which was obtained in the Utah study be taking the difference between Concept of Self and Ideal Self on the Index of Adjustment and Values. As Bills and other investigators have reported evidence to indicate that the Discrepancy Score is indicative of maladjustment, these rather consistent sex differences may be of considerable significance and probably should be explored further.

### Concept of Self Data

It will be noted in Table 10-25 that only two significant between-treatment differences occurred in the Sample VIII Concept of Self scores. During the ninth grade, District R average boys scored significantly higher than District A average boys. During the tenth grade, District A superior boys scored significantly higher than boys of comparable ability in District R.

Within-treatment comparisons revealed three significant differences. Only the difference favoring District R average boys over slow boys was significant for both years.

### Acceptance of Self Data

Comparison of Acceptance of Self means revealed only one statistically significant between-treatment difference for each of the two years in which the IAV was administered to this sample. District R superior girls were significantly higher during the ninth grade year, and District R average boys were significantly higher during the tenth grade year. Within-treatment differences show District A superior girls to have obtained a somewhat lower Acceptance of Self score than average or slow girls during both years with the former of these differences reaching statistical significance. In District R, only one within-treatment difference was significant, favoring average boys as compared with superior boys.

Table 10-24

Sex Differences in the IAV Scores for Sample VIII

| Groups | 59-60 IAV $M_{Boys}$ | 59-60 IAV $M_{Girls}$ | 59-60 IAV t | 60-61 IAV $M_{Boys}$ | 60-61 IAV $M_{Girls}$ | 60-61 IAV t |
|---|---|---|---|---|---|---|
| Concept of Self | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 86.98 | 83.94 | A.S. | 143.02 | 144.64 | N.S. |
| District A - Average Boys vs. District A - Average Girls | 82.69 | 84.08 | N.S. | 140.81 | 141.28 | N.S. |
| District A - Slow Boys vs. District A - Slow Girls | 86.38 | 85.31 | N.S. | 137.19 | 140.61 | N.S. |
| District R - Superior Boys vs. District R - Superior Girls | 85.86 | 86.35 | N.S. | 135.46 | 146.37 | 3.96 |
| District R - Average Boys vs. District R - Average Girls | 88.09 | 85.41 | N.S. | 143.60 | 143.86 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 83.80 | 83.06 | N.S. | 134.56 | 140.88 | N.S. |
| Acceptance of Self | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 88.61 | 78.79 | 3.56 | 139.05 | 130.80 | A.S. |
| District A - Average Boys vs. District A - Average Girls | 85.32 | 84.59 | N.S. | 134.72 | 136.27 | N.S. |
| District A - Slow Boys vs. District A - Slow Girls | 88.97 | 85.54 | N.S. | 139.67 | 139.00 | N.S. |
| District R - Superior Boys vs. District R - Superior Girls | 86.74 | 84.32 | N.S. | 134.00 | 133.46 | N.S. |
| District R - Average Boys vs. District R - Average Girls | 88.82 | 85.95 | N.S. | 145.04 | 138.14 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 86.24 | 88.03 | N.S. | 137.30 | 134.40 | N.S. |
| Ideal Self | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 101.36 | 101.13 | N.S. | 167.56 | 172.98 | 2.62 |
| District A - Average Boys vs. District A - Average Girls | 96.42 | 100.28 | 3.86 | 163.36 | 169.00 | 2.71 |
| District A - Slow Boys vs. District A - Slow Girls | 91.28 | 96.15 | N.S. | 157.14 | 166.56 | A.S. |
| District R - Superior Boys vs. District R - Superior Girls | 100.64 | 100.78 | N.S. | 164.16 | 171.31 | 3.33 |
| District R - Average Boys vs. District R - Average Girls | 98.70 | 100.95 | N.S. | 164.04 | 169.03 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 90.00 | 96.11 | 2.55 | 152.52 | 165.28 | 2.21 |
| Discrepancy Score | | | | | | |
| District A - Superior Boys vs. District A - Superior Girls | 14.38 | 18.33 | 2.16 | 28.74 | 32.26 | N.S. |
| District A - Average Boys vs. District A - Average Girls | 14.84 | 16.25 | N.S. | 29.36 | 30.73 | N.S. |
| District A - Slow Boys vs. District A - Slow Girls | 7.14 | 12.00 | 1.97 | 32.24 | 32.33 | N.S. |
| District R - Superior Boys vs. District R - Superior Girls | 14.80 | 14.63 | N.S. | 32.12 | 28.06 | A.S. |
| District R - Average Boys vs. District R - Average Girls | 11.03 | 16.17 | 2.34 | 25.32 | 29.24 | N.S. |
| District R - Slow Boys vs. District R - Slow Girls | 10.12 | 13.25 | N.S. | 31.59 | 28.20 | N.S. |

## Table 10-25

Sample VIII - Differences in Concept of Self Scores Over a Two-Year Period

| | 9th Grade | | | | 10th Grade | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Significant Differences | | | | Significant Differences | |
| Group | N | Mean Score | Between Treatments | Within Treatments | N | Mean Score | Between Treatments | Within Treatments |
| **District A** | | | | | | | | |
| 1. Superior Boys | 47 | 86.98 | | gr.3[a]<br>2.53 | 43 | 143.02 | gr.7<br>2.55 | |
| 2. Superior Girls | 63 | 83.94 | | | 61 | 144.64 | | |
| 3. Average Boys | 126 | 82.69 | | | 99 | 140.82 | | |
| 4. Average Girls | 102 | 84.08 | | | 86 | 141.27 | | |
| 5. Slow Boys | 29 | 86.38 | | | 21 | 137.19 | | |
| 6. Slow Girls | 26 | 85.31 | | | 18 | 140.61 | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 88 | 85.86 | | | 67 | 135.46 | | |
| 8. Superior Girls | 63 | 86.35 | | | 54 | 146.37 | | gr.12[b]<br>A.S. |
| 9. Average Boys | 33 | 88.09 | gr.3<br>2.79 | gr.11<br>2.10 | 25 | 143.60 | | gr.7,11<br>2.14;A.S. |
| 10. Average Girls | 41 | 85.41 | | | 29 | 143.86 | | |
| 11. Slow Boys | 41 | 83.80 | | | 27 | 134.56 | | |
| 12. Slow Girls | 36 | 83.06 | | | 24 | 140.88 | | |

[a] Indicates that this group(s) is significantly lower than group in left hand column; number indicates size of t.

[b] A.S. indicates approaching significance, - t=1.65 to 1.95.

## Table 10-26

Sample VIII - Differences in Acceptance of Self Scores Over a Two-Year Period

| | 9th Grade | | | | 10th Grade | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Significant Differences | | | | Significant Differences | |
| Group | N | Mean Score | Between Treatments | Within Treatments | N | Mean Score | Between Treatments | Within Treatments |
| **District A** | | | | | | | | |
| 1. Superior Boys | 47 | 88.62 | | | 43 | 139.05 | | |
| 2. Superior Girls | 63 | 78.79 | | | 61 | 130.80 | | |
| 3. Average Boys | 126 | 85.32 | | | 99 | 134.17 | | |
| 4. Average Girls | 102 | 84.59 | | gr.2<br>2.57 | 86 | 136.27 | | |
| 5. Slow Boys | 29 | 88.97 | | | 21 | 139.67 | | |
| 6. Slow Girls | 26 | 85.54 | | gr.2<br>A.S.[b] | 18 | 139.00 | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 88 | 86.74 | | | 67 | 134.00 | | |
| 8. Superior Girls | 63 | 84.32 | gr.2[a]<br>2.13 | | 54 | 133.46 | | |
| 9. Average Boys | 33 | 88.82 | | | 25 | 145.04 | gr.3<br>2.00 | gr.7,11<br>2.52;A.S. |
| 10. Average Girls | 41 | 85.95 | | | 29 | 138.14 | | |
| 11. Slow Boys | 41 | 86.24 | | | 27 | 137.30 | | |
| 12. Slow Girls | 36 | 88.03 | | | 24 | 134.40 | | |

---

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

Table 10-27

Sample VIII - Differences in Discrepancy Scores Over a Two-Year Period

| | 9th Grade Significant Differences | | | | 10th Grade Significant Differences | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score | Between Treatments | Within Treatments | N | Mean Score | Between Treatments | Within Treatments |
| **District A** | | | | | | | | |
| 1. Superior Boys | 47 | 14.38 | | gr.5<br>3.90 | 43 | 28.74 | | |
| 2. Superior Girls | 63 | 18.33 | gr.8<br>2.18[a] | gr.6<br>2.59 | 61 | 32.26 | | |
| 3. Average Boys | 126 | 14.84 | gr.9[b]<br>A.S. | gr.5<br>3.71 | 99 | 29.36 | | |
| 4. Average Girls | 102 | 16.25 | | gr.6<br>A.S. | 86 | 30.73 | | |
| 5. Slow Boys | 29 | 7.14 | | | 21 | 32.24 | | |
| 6. Slow Girls | 26 | 12.00 | | | 18 | 32.33 | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 88 | 14.81 | | gr.9,11<br>2.03;2.16 | 67 | 32.12 | | gr.9<br>2.31 |
| 8. Superior Girls | 63 | 14.63 | | | 54 | 28.06 | | |
| 9. Average Boys | 33 | 11.03 | | | 25 | 25.32 | | |
| 10. Average Girls | 41 | 16.17 | | | 29 | 29.24 | | |
| 11. Slow Boys | 41 | 10.12 | | | 27 | 31.59 | | |
| 12. Slow Girls | 36 | 13.25 | | | 24 | 28.20 | | |

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

Table 10-28

Sample VIII - Differences in Ideal Self Scores Over a Two-Year Period

| | 9th Grade | | Significant Differences | | 10th Grade | | Significant Differences | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score | Between Treatments | Within Treatments | N | Mean Score | Between Treatments | Within Treatments |
| District A | | | | | | | | |
| 1. Superior Boys | 47 | 101.36 | | gr.3,5<br>3.61[a];4.34 | 43 | 167.56 | | gr.3,5<br>A.S.;2.72 |
| 2. Superior Girls | 63 | 101.13 | | gr.6<br>2.67 | 61 | 172.98 | | gr.4,6<br>1.98;2.29 |
| 3. Average Boys | 126 | 96.43 | | gr.5<br>2.39 | 99 | 163.36 | | gr.5[b]<br>A.S. |
| 4. Average Girls | 102 | 100.28 | | gr.6<br>2.85 | 86 | 169.00 | | |
| 5. Slow Boys | 29 | 91.28 | | | 21 | 157.14 | | |
| 6. Slow Girls | 26 | 96.15 | | | 18 | 166.56 | | |
| District R | | | | | | | | |
| 7. Superior Boys | 88 | 100.64 | | gr.11<br>7.02 | 67 | 164.16 | | gr.11<br>2.79 |
| 8. Superior Girls | 63 | 100.78 | | gr.12<br>2.97 | 54 | 171.31 | | gr.12<br>2.56 |
| 9. Average Boys | 33 | 98.70 | | gr.11<br>4.08 | 25 | 164.04 | | gr.11<br>A.S. |
| 10. Average Girls | 41 | 100.95 | | gr.12<br>2.53 | 29 | 169.03 | | |
| 11. Slow Boys | 41 | 90.00 | | | 27 | 152.52 | | |
| 12. Slow Girls | 36 | 96.11 | | | 24 | 165.28 | | |

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

### Discrepancy Score Data

Only one significant between-treatment difference on Discrepancy Score means was obtained for this sample. District A superior girls had a higher mean than comparable District R girls. There appeared to be some tendency towards a positive relationship between ability level and Discrepancy Scores, thus suggesting poorer adjustment among the more able pupils. This was reflected in the within-treatment differences for both Districts A and R on the ninth grade testing. The presence of only one significant difference in the tenth grade replication, however, suggests that at the high school level the somewhat changed situation brought about by the wide range of electives available in both districts and less distinct ability groupings in District A, may have reduced the importance of the grouping treatments in terms of their influence upon pupils' self-concept.

### Ideal Self Data

No significant between-treatment differences were found on the Ideal Self scores. Most of the within-treatment differences in both Districts, however, were statistically significant on this score, with the differences generally favoring pupils of higher ability. Nine of the 12 within-treatment differences for District A were statistically significant with two of the three remaining comparisons approaching significance. The within-treatment differences were somewhat less marked for District R but followed the same general pattern with six reaching statistical significance and one approaching significance.

## SAMPLE IX RESULTS

### Introduction

The Index of Adjustment and Values was administered to Sample IX pupils during their ninth and eleventh grade years. Analysis procedures for this sample followed the same pattern employed with previous samples excepting that the groups were not divided by sex for analysis of the eleventh grade data. The reader will recall that during the high school years, both districts required students to take only one year of mathematics and science. At the eleventh grade level, many pupils in Sample IX were not taking courses in these areas and therefore could not be classified in terms of ability for that school year. This led to considerable reduction in the sample and reduced the size of some groups to a point where division by sex would have produced groups too small for analysis. Even without subdividing, by sex, some of the eleventh grade groups are smaller than is desirable and the results based on these data cannot be accepted with the same degree of confidence as the IAV data for other samples.

### Sex Differences

Table 10-29 summarizes sex differences for Sample IX on the ninth grade administration of the IAV. In reviewing the significant sex differences, it is interesting to note that in District A, superior girls obtained significantly more favorable scores on Concept of Self, Acceptance of Self, and Discrepancy Scores than did comparable boys in this district. Three of the six sex differences on the Ideal Self score were statistically significant with girls obtaining the higher mean in each case, following the pattern set by other samples. Girls have generally obtained higher Discrepancy Scores than boys, in the comparisons of other samples. For Sample IX, however, although slow girls received a significantly higher score than slow boys, the reverse was true of the mean scores for District A superior pupils.

### Concept of Self Data

Between and within-treatment comparisons of Concept of Self scores for Sample IX are found in Table 10-30. Ninth grade data revealed only one significant between-treatment difference; average District R girls scoring higher than comparable girls in District A. In District A, within-treatment comparisons showed the superior girls to have obtained significantly higher self-concept scores than either average or slow girls. In District R superior boys received higher scores than average boys while average girls received higher scores than average girls. No significant differences were found in comparing subgroups on the eleventh grade data.

### Acceptance of Self Data

No statistically significant differences were found in either the between-treatment or within-treatment comparisons of Sample IX Acceptance of Self scores (see Table 10-31).

### Discrepancy Score Data

Between-treatment comparisons on the Discrepancy Score showed a consistent tendency for District A pupils to obtain higher (less favorable) scores than comparable pupils in District R. Only the difference between superior girls reached statistical significance but it will be noted in Table 10-32 that three other differences approached significance. Three within-treatment differences were significant for the ninth grade data, two in District A and one in District R. These significant differences generally reflected a tendency for girls of lower ability level to have higher (less favorable) Discrepancy Scores.

### Ideal Self Data

It will be noted in Table 10-33 that only one significant between-treatment difference was found in Ideal Self; superior pupils in District A scoring higher than those in District R. The ninth grade data reflected a fairly consistent pattern of within-treatment differences generally favoring the group of higher ability level.

### Discussion

In reviewing the data for Sample IX, we find that the ninth grade results are generally in line with those found in other samples. The almost complete lack of significant differences between IAV scores of different subgroups at the eleventh grade level is probably caused by the small number of cases and the fact that the actual differences between the ability grouping and random treatments is much smaller at this level than in the junior high school and elementary grades.

## Table 10-29

### Sex Differences for Sample IX

| Groups | $M_{Boys}$ | $M_{Girls}$ | t |
|---|---|---|---|
| Concept of Self | | | |
| District A Superior Boys vs. District A Superior Girls | 137.00 | 146.89 | 2.50 |
| District A Average Boys vs. District A Average Girls | 136.89 | 138.62 | N.S. |
| District A Slow Boys vs. District A Slow Girls | 130.49 | 131.93 | N.S. |
| District R Superior Boys vs. District R Superior Girls | 143.52 | 141.20 | N.S. |
| District R Average Boys vs. District R Average Girls | 134.05 | 146.48 | 3.42 |
| District R Slow Boys vs. District R Slow Girls | 137.25 | 136.67 | N.S. |
| Acceptance of Self | | | |
| District A Superior Boys vs. District A Superior Girls | 132.61 | 146.26 | 2.11 |
| District A Average Boys vs. District A Average Girls | 137.68 | 136.62 | N.S. |
| District A Slow Boys vs. District A Slow Girls | 135.44 | 136.79 | N.S. |
| District R Superior Boys vs. District R Superior Girls | 142.48 | 137.84 | N.S. |
| District R Average Boys vs. District R Average Girls | 136.89 | 142.70 | N.S. |
| District R Slow Boys vs. District R Slow Girls | 141.84 | 139.73 | N.S. |
| Ideal Self Scores | | | |
| District A Superior Boys vs. District A Superior Girls | 166.67 | 170.40 | N.S. |
| District A Average Boys vs. District A Average Girls | 158.39 | 166.82 | 2.92 |
| District A Slow Boys vs. District A Slow Girls | 150.90 | 164.29 | 2.34 |
| District R Superior Boys vs. District R Superior Girls | 168.71 | 168.84 | N.S. |
| District R Average Boys vs. District R Average Girls | 153.32 | 167.55 | 2.55 |
| District R Slow Boys vs. District R Slow Girls | 152.41 | 155.73 | N.S. |
| Discrepancy Scores | | | |
| District A Superior Boys vs. District A Superior Girls | 34.39 | 26.14 | 2.69 |
| District A Average Boys vs. District A Average Girls | 29.50 | 33.47 | N.S. |
| District A Slow Boys vs. District A Slow Girls | 32.56 | 39.29 | N.S. |
| District R Superior Boys vs. District R Superior Girls | 28.71 | 31.67 | N.S. |
| District R Average Boys vs. District R Average Girls | 29.43 | 27.27 | N.S. |
| District R Slow Boys vs. District R Slow Girls | 25.38 | 34.97 | 2.39 |

Table 10-30

Sample IX - Differences in Concept of Self Scores Over a Two-Year Period

| | 9th Grade Significant Differences | | | | 11th Grade Significant Differences[c] | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| **District A** | | | | | | | | |
| 1. Superior Boys | 18 | 137.00 | | | 1. 27 | 145.33 | | |
| 2. Superior Girls | 35 | 146.88 | gr.8[b]<br>A.S. | gr.4,6[a]<br>t=2.36,3.27 | | | | |
| 3. Average Boys | 28 | 136.89 | | | 2. 73 | 142.66 | | |
| 4. Average Girls | 34 | 138.62 | | | | | | |
| 5. Slow Boys | 39 | 130.49 | | | 3. 13 | 147.08 | | |
| 6. Slow Girls | 14 | 131.93 | | | | | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 21 | 143.52 | | gr.9<br>t=2.11 | 4. 75 | 144.13 | | |
| 8. Superior Girls | 45 | 141.20 | | | | | | |
| 9. Average Boys | 37 | 134.05 | | | 5. 30 | 145.43 | | |
| 10. Average Girls | 33 | 146.48 | gr.4<br>t=2.14 | gr.12<br>t=2.35 | | | | |
| 11. Slow Boys | 32 | 137.25 | | | 6. 30 | 142.67 | | |
| 12. Slow Girls | 30 | 136.67 | | | | | | |

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

[c] Sufficient cases were not available to analyze sexes separately.

## Table 10-31

Sample IX - Differences in Acceptance of Self Scores Over a Two-Year Period

| Group | 9th Grade Significant Differences N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | 11th Grade Significant Differences[b] N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
|---|---|---|---|---|---|---|---|---|
| **District A** | | | | | | | | |
| 1. Superior Boys | 18 | 132.61 | | | 1. 27 | 143.15 | | |
| 2. Superior Girls | 35 | 146.26 | gr.8[a]<br>A.S. | gr.4<br>A.S. | | | | |
| 3. Average Boys | 28 | 137.68 | | | 2. 73 | 137.64 | | |
| 4. Average Girls | 34 | 136.62 | | | | | | |
| 5. Slow Boys | 39 | 135.44 | | | 3. 13 | 141.69 | | |
| 6. Slow Girls | 14 | 136.79 | | | | | | |
| **District R** | | | | | | | | |
| 7. Superior Boys | 21 | 142.48 | | | 4. 75 | 138.97 | | |
| 8. Superior Girls | 45 | 137.84 | | | | | | |
| 9. Average Boys | 37 | 136.89 | | | 5. 30 | 145.47 | | |
| 10. Average Girls | 33 | 142.70 | | | | | | |
| 11. Slow Boys | 32 | 141.84 | | | 6. 30 | 143.50 | | |
| 12. Slow Girls | 30 | 139.73 | | | | | | |

[a] A.S. indicates approaching significance - t=1.65 through 1.95.

[b] Sufficient cases were not available to analyze sexes separately.

Table 10-32

Sample IX - Differences in Discrepancy Scores Over a Two-Year Period

| | 9th Grade Significant Differences | | | | 11th Grade Significant Differences[c] | | | |
|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment | N | Mean Score (JHS - IAV) | Between Treatment | Within Treatment |
| District A | | | | | | | | |
| 1. Superior Boys | 18 | 34.39 | gr.7<br>A.S. | gr.3[b]<br>A.S. | 1. 27 | 26.15 | | |
| 2. Superior Girls | 35 | 26.14 | gr.8[a]<br>1.96 | | | | | |
| 3. Average Boys | 28 | 29.50 | | | 2. 73 | 29.82 | | |
| 4. Average Girls | 34 | 33.47 | gr.10<br>A.S. | gr.2<br>2.26 | | | | |
| 5. Slow Boys | 39 | 32.56 | gr.11<br>A.S. | | 3. 13 | 25.08 | | |
| 6. Slow Girls | 14 | 39.29 | | gr.2<br>3.48 | | | | |
| District R | | | | | | | | |
| 7. Superior Boys | 21 | 28.71 | | | 4. 75 | 27.80 | | |
| 8. Superior Girls | 45 | 31.67 | | | | | | |
| 9. Average Boys | 37 | 29.43 | | | 5. 30 | 26.70 | | |
| 10. Average Girls | 33 | 27.27 | | | | | | |
| 11. Slow Boys | 32 | 25.38 | | | 6. 30 | 32.83 | | |
| 12. Slow Girls | 30 | 34.97 | | gr.10<br>2.11 | | | | |

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

[c] Sufficient cases were not available to analyze sexes separately.

## Table 10-33

### Sample IX - Differences in Ideal Self Scores Over a Two-Year Period

| | | 9th Grade Significant Differences | | | | | 11th Grade Significant Differences[c] | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | N | Mean Score (JHS - IAV) | Between Treatments | Within Treatments | | N | Mean Score (JHS - IAV) | Between Treatments | Within Treatments |
| **District A** | | | | | | | | | |
| 1. Superior Boys | 18 | 166.67 | | gr.3,5<br>2.19[a];3.26 | 1. | 27 | 172.93 | gr.4<br>2.56 | gr.2,3<br>2.47;A.S. |
| 2. Superior Girls | 35 | 170.40 | | gr.6<br>A.S.[b] | | | | | |
| 3. Average Boys | 28 | 158.39 | | gr.5<br>A.S. | 2. | 73 | 164.99 | | |
| 4. Average Girls | 34 | 166.82 | | | | | | | |
| 5. Slow Boys | 39 | 150.90 | | | 3. | 13 | 166.85 | | |
| 6. Slow Girls | 14 | 164.29 | | | | | | | |
| **District R** | | | | | | | | | |
| 7. Superior Boys | 21 | 168.71 | | gr.9,11<br>2.28;4.03 | 4. | 75 | 164.61 | | |
| 8. Superior Girls | 45 | 168.84 | | gr.12<br>3.73 | | | | | |
| 9. Average Boys | 37 | 153.32 | | | 5. | 30 | 168.23 | | |
| 10. Average Girls | 33 | 167.55 | | gr.12<br>2.89 | | | | | |
| 11. Slow Boys | 32 | 152.41 | | | 6. | 30 | 163.43 | | |
| 12. Slow Girls | 30 | 155.73 | | | | | | | |

---

[a] Indicates that this group is significantly lower than group in left hand column, number indicates size of t.

[b] A.S. indicates approaching significance - t=1.65 through 1.95.

[c] Sufficient cases were not available to analyze sexes separately.

## RESULTS ON OTHER SELF-CONCEPT MEASURES

### Sense of Personal Worth

The Sense of Personal Worth scale from the California Test of Personality was administered to Sample IV pupils during their fifth grade year. This measure appears to be related to the Acceptance of Self score on the Index of Adjustment and Values. At this time, District A pupils had been exposed to the ability grouping situation for about one and one-half years. Results for this sample may be found in Table 12-5 in Chapter 12. All six District R samples scored higher than District A pupils of comparable ability on this measure. These differences were significant at the .05 level for both superior boys and superior girls. District R average boys were significantly higher than comparable District A boys at the .01 level, but the difference between the average girls in both districts did not quite reach statistical significance. The differences were not significant for slow pupils in the two treatments.

Within-treatment differences on this scale reflect about the same trend in both districts. Superior pupils obtained the highest scores with the average and slow pupils following in that order. In District A, three of the six within-treatment differences were statistically significant, while four differences were found to be significant for District R (see Table 12-6 in Chapter 12). There was also a consistent pattern of sex difference in the two districts with girls obtaining higher scores on this measure than boys. All sex differences in District A were statistically significant, and the differences favoring superior and slow girls were significant in District R.

### Self-Acceptance

The Self-Acceptance Scale of the California Psychological Inventory was administered to Sample VI during their eight grade year, and to Sample VII during their ninth grade year. The items contained in this scale, and the concurrent validity data reported in the test manual indicate that this measure is related to the Acceptance of Self score on the IAV. Results for these two samples may be found in Tables 12-12 and 12-13 in Chapter 12.

In comparing Sample VI subgroups in the two treatments it was found that District R pupils received more favorable Self-Acceptance scores in every case than comparable pupils of District A. These between-treatment differences were statistically significant at the .01 level for both average boys and average girls and approached significance for slow girls. Between-treatment comparisons of the Self-Acceptance Scale results for Sample VII again revealed a tendency for District R pupils to have higher scores although differences were smaller than found in Sample VI, only one, between average girls, was statistically significant.

Within-treatment differences in Sample VI were generally significant in both districts, with pupils of higher ability obtaining higher scores. In District A, four of the six differences between ability levels were statistically significant, as compared with District R where only two of the within-treatment differences were significant. Within-treatment results for pupils in Sample VII closely corresponded to the Sample VI data with ability level paralleling self acceptance. These differences were again somewhat more pronounced in District A, where four of the six differences were statistically significant, and the remaining two approached significance. In District R, three of these differences reached statistical significance and one of the remaining three approached significance.

There were no statistically significant sex differences in either district between Sample VI and VII boys and girls of comparable ability.

## About Myself Scale

The SRA Junior Inventory was administered to Sample IV pupils during their fifth grade year. Between-treatment differences for the four samples who were administered the About Myself Scale may be found in Table 10-34. Although five of the six between-treatment comparisons showed higher (less favorable) scores for District A, it will be noted that the only significant differences found on this scale were between average pupils in the two districts with District A average boys and girls both receiving significantly higher scores than comparable groups in District R. As the pupil's score indicates the number of problems he is having in a given area, high scores are undesirable in this measure.

The SRA Youth Inventory was administered to Sample VI during their seventh grade year. Between-treatment comparisons showed no significant differences, although District R pupils checked a slightly larger mean number of problems in this area. Within-treatment comparisons showed a consistent trend for pupils of higher ability to check fewer problems although these differences were generally small. Superior girls in both districts checked significantly fewer problems than slow girls, but no other significant within-treatment differences emerged. In the six sex differences, boys received slightly lower (more favorable) mean scores than girls in every case. Only one of these differences, however, between average boys and girls in District A reached statistical significance.

The About Myself Scale was also administered to Sample VII pupils during their eighth grade year. It will be noted in Table 10-34 that differences between mean scores for this sample were small although generally favoring District R pupils.

Comparisons between pupils at different ability levels again revealed a consistent tendency for pupils of higher ability to check fewer problems on this scale. Two of the six within-treatment differences for the District A sample and three of the differences for District R sample reached statistical significance.

A comparison of boys' and girls' subgroups in the two districts showed a consistent tendency for girls to check a larger number of problems than boys. In District A, two of the sex differences were significant beyond the .01 level, and the third approached significance. These differences were somewhat smaller in District R with one comparison, that between average boys and girls reaching statistical significance at the .05 level.

The About Myself Scale was administered to Sample VIII pupils during their ninth grade year. In this sample, District R pupils received higher (less favorable) scores on this measure in all six of the between-treatment comparisons. The difference between the mean scores of slow boys was statistically significant at the .01 level, while differences between superior boys and average girls in the two treatments approached significance. Within-treatment differences for Sample VIII did not reflect quite as consistent a tendency for better pupils to check fewer problems as was found in the Sample VII data. In District A, superior and slow boys checked fewer problems than average boys. For District A girls, the superior group checked fewer problems than the average or slow groups, the latter two being almost identical. The within-treatment differences followed the same pattern as found in Sample VII, however, for District R pupils with the number of problems checked being inversely related to

Table 10-34

Between-Treatment Differences on the SRA About Myself Scale

| | Sample IV<br>5th Grade<br>SRA Jr. Inv. | | | Sample VI<br>7th Grade<br>SRA Youth Inv. | | | Sample VII<br>8th Grade<br>SRA Youth Inv. | | | Sample VIII<br>9th Grade<br>SRA Youth Inv. | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys<br>District A vs. R | 22.40 | 23.03 | NS[a] | 8.70 | 10.19 | NS | 7.84 | 9.21 | NS | 7.28 | 8.98 | NS |
| Superior Girls<br>District A vs. R | 25.32 | 22.06 | NS | 9.81 | 10.76 | NS | 12.03 | 10.07 | NS | 10.15 | 11.14 | NS |
| Average Boys<br>District A vs. R | 29.78 | 23.97 | 2.36 | 9.01 | 10.13 | NS | 8.49 | 10.38 | AS[b] | 7.57 | 11.04 | AS |
| Average Girls<br>District A vs. R | 32.55 | 25.51 | 2.99 | 11.45 | 12.19 | NS | 13.06 | 14.10 | NS | 9.19 | 10.77 | NS |
| Slow Boys<br>District A vs. R | 31.93 | 29.05 | NS | 10.94 | 12.61 | NS | 11.31 | 12.60 | NS | 10.68 | 13.98 | AS |
| Slow Girls<br>District A vs. R | 32.41 | 31.82 | NS | 13.68 | 14.65 | NS | 16.06 | 14.35 | NS | 12.15 | 14.14 | NS |

[a] Not Significant

[b] Approaching significance, t from 1.65 to 1.95 inclusive.

ability level. One of these within-treatment differences in District R was statistically significant and two other approached significance; in all cases favoring the group of higher ability.

No significant sex differences were found in comparing boys and girls of comparable ability within the two treatments. In all six of the sex comparisons, boys checked fewer problems than girls with these differences approaching significance for slow pupils in District A and for average pupils in District R.

## SUMMARY AND CONCLUSIONS

This section will attempt to pull together the rather extensive data relative to self-concept collected in the Utah study and draw conclusions that appear to be justified.

### Between-treatment Differences

Between-treatment differences on the four scores obtained from Bills' Index of Adjustment and Values are summarized in Table 10-35. For each of the four IAV variables, 66 between-treatment comparisons were made, 22 at each ability level. The largest number of between-treatment comparisons, 18 were made for Sample VI pupils. Fifteen comparisons were made for Sample IV, 12 for Samples VII and VIII, and nine for Sample IX. Data collected from the sixth through the ninth grade levels probably provide the most meaningful results on self-concept. Differences between fifth grade pupils were based on the elementary school form of the IAV which is shorter and somewhat less reliable than the junior high school form. Differences between pupils in the high school are less meaningful because of the less clearcut distinction between the two grouping treatments at that level. Except for data collected on fifth grade pupils in Sample IV, and eleventh grade pupils in Sample IX, boys and girls were placed in separate groups for the IAV analyses.

#### Concept of Self

Of the 66 between-treatment comparisons on the Concept of Self score, 23 reached or approached statistical significance. All but two of these 23 differences were in a direction favoring District R pupils. Eight of these differences were between superior pupils in the two districts, 10 between average pupils, and five between slow pupils. It is interesting to note in the Concept of Self analysis that 15 of the 23 differences reaching or approaching statistical significance were between girls' groups while only seven of these differences were between boys' groups. One of the significant differences, for Sample IV, was based on groups not divided by sex.

The About Myself Scale from the SRA Youth Inventory also appears to measure self-concept. The two significant between-treatment differences obtained on this scale favor District R pupils, thus agreeing with the pattern that emerged from analysis of the Index of Adjustment and Values.

The extent and consistency of the Concept of Self data seem to justify the conclusion that pupils in random grouped classrooms tend to obtain more favorable Concept of Self scores than comparable pupils in ability grouped classrooms, at all ability levels. The data also appear to support the conclusion that the two grouping treatments had a somewhat greater effect upon girls than boys. These data may

Table 10-35

Summary of Significant Between-Treatment Differences on the IAV

| | Sample IV Grade | | | Sample VI Grade | | | Sample VII Grade | | Sample VIII Grade | | Sample IX Grade | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5[a] | 6 | 7 | 7 | 8 | 9 | 8 | 9 | 9 | 10 | 9 | 11[a] |
| **Concept of Self** | | | | | | | | | | | | |
| Superior Boys | R[b] | R | R | | | | | | | A | | |
| Superior Girls | | R* | | R | | | R | | | | A* | |
| Average Boys | | R | R | | | | | | R | | | |
| Average Girls | | R | R | R | R | R | | R | | | R | |
| Slow Boys | | | R* | | | | | | | | | |
| Slow Girls | | R* | R | | | | R | R | | | | |
| **Acceptance of Self** | | | | | | | | | | | | |
| Superior Boys | R | R | R | | A* | A* | | | | | | |
| Superior Girls | | | | R* | | | R | R | R | | A* | |
| Average Boys | R | R | R | | | | | | | R | | |
| Average Girls | | | | R | | R* | | R | | | | |
| Slow Boys | R | | R | | | | | | | | | |
| Slow Girls | | | | R* | R | | R* | | | | | |
| **Ideal Self** | | | | | | | | | | | | |
| Superior Boys | | | | | | | | | | | | A |
| Superior Girls | | | | R | | | | | | | | |
| Average Boys | | R | R | | | | | | | | | |
| Average Girls | | R | | R | | | | R* | | | | |
| Slow Boys | | | | R | | | | | | | | |
| Slow Girls | | | | | | | R | | | | | |
| **Discrepancy Score** | | | | | | | | | | | | |
| Superior Boys | A | A | A | A* | | A* | | | | | A* | |
| Superior Girls | | | A* | | | | A* | | A | | R | |
| Average Boys | A | A | A | | | | | | A* | | | |
| Average Girls | | | A | | A | A | | | | | A* | |
| Slow Boys | | | | | | | | | | | A* | |
| Slow Girls | | A | A | | | A* | | | | | | |

---

* Approached significance, (t value 1.65 to 1.95 inclusive).

[a] As data were not analysed separately by sex for this sample, a significant difference when found is indicated for both boys and girls.

[b] An "R" indicates that the District R subgroup had a significantly higher mean score; A indicates a significantly higher mean for the District A group.

[c] High Discrepancy Scores are considered undesirable.

reflect the greater preoccupation of girls with achievement and academic success and generally suggest that placement in an ability group tends to adversely affect girls' self-concept. This adverse effect is of least importance for pupils at the lower ability level.

## Acceptance of Self

Of the 66 between-treatment comparisons of IAV Acceptance of Self mean scores, 15 were statistically significant and seven others approached significance. These results reflected essentially the same trend found in analysis of the Concept of Self data. Six of the significant differences were between superior pupils, six between average pupils, and three between slow pupils in the two districts. All 15 of these significant between-treatment differences favored District R as did four of the seven differences that approached significance. It will be noted that the significance differences occurred between girls groups and boys groups with about equal frequency.

These results were generally supported by the Sense of Personal Worth scale from the California Test of Personality. All six between-treatment comparisons on this scale favored District R pupils with three being statistically significant.

The Self-Acceptance Scale of the California Psychological Inventory also yielded results in agreement with the IAV Acceptance of Self score. For Sample VI subgroups, District R pupils generally received more favorable scores on this scale than comparable pupils in District A with two of the between-treatment differences being statistically significant. Sample VII results on this scale reflected a similar pattern of superiority for District R pupils although only one of the six between-treatment differences was statistically significant.

When we consider all results relative to acceptance of self obtained in this study, a consistent pattern favoring District R pupils emerges. This pattern is most pronounced for superior and average pupils but is present to a lesser degree for the slower pupils. Based on these data it must be concluded that random grouping is consistently related to higher self acceptance for pupils at all ability levels and over most of the grade levels covered in this project.

## Discrepancy Score

The overall summary of between-treatment comparisons on the IAV Discrepancy Score shows 14 of the 66 differences to be statistically significant. Six of these differences were between superior pupils in the two districts. Six were between average pupils and two were between the slow pupils. All but one of the 14 significant differences showed higher (less favorable) Discrepancy Score means for District A subgroups. In addition to these significant differences, eight differences approached statistical significance, and all of these showed higher mean Discrepancy Scores for District A groups. Significant differences occurred with about equal frequency for boys' and girls' groups. It will be noted that these data follow the same general pattern that emerged in the overall results of the Concept of Self and Acceptance of Self analyses. It may be concluded that the two grouping treatments had a differential effect upon the pupils' Discrepancy Scores with ability grouping generally associated with less favorable scores. This pattern is most pronounced in the differences between the mean scores of superior pupils in the two treatments closely followed by average pupils and is substantially less marked in the differences between slow pupils. It may also be concluded that ability grouping appears to affect the Discrepancy Scores of girls' groups more unfavorable than boy's groups.

### Ideal Self

Of the 66 between-treatment comparisons of mean Ideal Self scores on the Index of Adjustment and Values, eight reached statistical significance. Seven of these eight differences favored District R pupils, with two occurring between superior groups, four between average groups and two between slow groups. It will be noted in Table 10-35 that the few significant differences on the Ideal Self scores are generally scattered. Only in the comparisons between average girls do we find significant differences occurring in more than one sample. From these data we may conclude that a slight trend favoring District R pupils is in evidence on the Ideal Self scores. This trend, producing relatively few significant results, suggests that the grouping system has little or no influence upon the Ideal Self scores.

## Within-treatment Differences

Comparisons between like-sex pupils of different ability levels in the two grouping treatments are summarized in Table 10-36. A total of 60 within-treatment comparisons are shown for each district on each of the IAV scores. Within-treatment comparisons for pupils in Sample IV, during the fifth grade, and Sample IX pupils during the eleventh grade were not included in the summary table. In Table 10-36 an entry of "1" indicates that a significant difference in mean scores was found favoring the group of higher ability, while a "2 " indicates that a significant difference favoring the group of lower ability was found in the sample and grade level indicated.

It will be noted that of the 60 comparisons between District A groups on Concept of Self means, 16 reached or approached statistical significance, all favoring the group of higher ability. The District R within-treatment comparisons yielded 15 differences that reached or approached statistical significance. All but one of these differences favored the group of higher ability. Thus, it may be concluded that there is a consistent tendency in both treatments for pupils of higher ability to have higher Concept of Self scores. This trend is no more pronounced in one grouping treatment than the other.

Of the total of 120 within-treatment comparisons on the Acceptance of Self score, we find only four that were statistically significant at or beyond the .05 level. Differences reaching or approaching significance are about equally divided between those favoring groups of higher and lower ability. The lack of a consistent trend and the small number of significant differences on the Acceptance of Self score leads to the conclusion that the pupil's ability level is unrelated to his Acceptance of Self score in both of the grouping treatments.

The within-treatment differences between Discrepancy Score means present a somewhat different picture in the two districts. In District A there were a total of 20 within-treatment differences that reached or approached statistical significance. Of these, 14 showed the group of higher ability to have higher (less favorable) Discrepancy Scores, while six showed the group of lower ability to have less favorable scores. For District A boys there appeared to be a tendency for superior pupils to have less favorable Discrepancy Scores while for District A girls the opposite trend was present.

Table 10-36

Summary of Significant Within Treatment Differences on the IAV

| | Sample IV | | Sample VI | | | Sample VII | | Sample VIII | | Sample IX |
|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 7 | 7 | 8 | 9 | 8 | 9 | 9 | 10 | 9 |
| **Concept of Self** | | | | | | | | | | |
| **District A** | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | | | | | | | | 1* | | 1 |
| 1. Superior Boys vs.<br>2. Slow Boys | | | | | | 1* | 1 | | | |
| 1. Average Boys vs.<br>2. Slow Boys | | | | | | | 1* | | | |
| 1. Superior Girls vs.<br>2. Average Girls | | | 1 | 1* | 1 | 1 | 1 | | | 1 |
| 1. Superior Girls vs.<br>2. Slow Girls | | | | | 1 | 1 | 1 | | | 1 |
| 1. Average Girls vs.<br>2. Slow Girls | | | | | | | | | | 1 |
| **District R** | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | 1[a] | | | | | 1* | | | 2 | |
| 1. Superior Boys vs.<br>2. Slow Boys | 1 | 1 | 1* | | | 1 | 1 | | | |
| 1. Average Boys vs.<br>2. Slow Boys | | 1 | | | | | | 1 | 1* | |
| 1. Superior Girls vs.<br>2. Average Girls | | | | | | 1 | | | | |
| 1. Superior Girls vs.<br>2. Average Girls | | | | | | 1* | | | 1* | |
| 1. Average Girls vs.<br>2. Slow Girls | | | 1* | | | | | | | |
| **Acceptance of Self** | | | | | | | | | | |
| **District A** | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Slow Boys | | | | | | | | | | |
| 1. Average Boys vs.<br>2. Slow Boys | | | | | | | | | | |
| 1. Superior Girls vs.<br>2. Average Girls | | | | | | | | 2 | | 1* |
| 1. Superior Girls vs.<br>2. Slow Girls | | | | | | | | 2* | | |
| 1. Average Girls vs.<br>2. Slow Girls | | | | | | | | | | |

| | Sample IV | | Sample VI | | | Sample VII | | Sample VIII | | Sample IX |
|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 7 | 7 | 8 | 9 | 8 | 9 | 9 | 10 | 9 |
| District R | | | | | | | | | | |
| 1. Superior Boys vs. | | | | | | | | | | |
| 2. Average Boys | | | | | 2 | 1* | | 2 | | |
| 1. Superior Boys vs. | | | | | | | | | | |
| 2. Slow Boys | | | | | 2* | | | | | |
| 1. Average Boys vs. | | | | | | | | | | |
| 2. Slow Boys | 1 | | | | | | | 1* | | |
| 1. Superior Girls vs. | | | | | | | | | | |
| 2. Average Girls | | | | | | | | | | |
| 1. Superior Girls vs. | | | | | | | | | | |
| 2. Slow Girls | | | | | | | | | | |
| 1. Average Girls vs. | | | | | | | | | | |
| 2. Slow Girls | | | | | | | | | | |

* These differences approached statistical significance (t=1.65-1.95)

[a] A "1" indicates that group 1 of the two groups being compared made a significantly higher mean score.

## Table 10-37

### Summary of Significant Within Treatment Differences on the IAV (Con.)

| | Sample IV | | Sample VI | | | Sample VII | | Sample VIII | | Sample IX |
|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 7 | 7 | 8 | 9 | 8 | 9 | 9 | 10 | 9 |
| Discrepancy Scores | | | | | | | | | | |
| District A | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | 1 | | 1 | | | | | | | 1* |
| 1. Superior Boys vs,<br>2. Slow Boys | 1 | | 1 | | | 1 | | 1 | | |
| 1. Average Boys vs.<br>2. Slow Boys | | | | | | 1* | | 1 | | |
| 1. Superior Girls vs.<br>2. Average Girls | | 2 | 1* | | 2 | | | 1* | | 2 |
| 1. Superior Girls vs.<br>2. Slow Girls | 1 | 2 | | | 2 | | | 1 | | 2 |
| 1. Average Girls vs.<br>2, Slow Girls | 1 | | | | | | | | | |
| District R | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | | | | | | 1 | 1* | 1 | 1 | |
| 1. Superior Boys vs.<br>2. Slow Boys | | | | | | | | 1 | | |
| 1. Average Boys vs.<br>2. Slow Boys | | | | | | | | | | |
| 1. Superior Girls vs.<br>2. Average Girls | | | | | | | | | | |
| 1. Superior Girls vs.<br>2. Slow Girls | | | | | | | | | | |
| 1. Average Girls vs.<br>2. Slow Girls | | | | | | | | | | 2 |
| Ideal Self | | | | | | | | | | |
| District A | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | 1 | 1 | 1 | 1* | | 1 | 1* | 1 | 1* | 1 |
| 1. Superior Boys vs.<br>2. Slow Boys | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1. Average Boys vs.<br>2. Slow Boys | | 1 | | | | 1 | 1 | 1 | 1* | 1* |
| 1. Superior Girls vs.<br>2. Average Girls | 1 | | 1 | | 1* | 1 | 1 | | 1 | |
| 1. Superior Girls vs.<br>2. Slow Girls | 1 | | 1 | 1 | | 1 | 1 | 1 | 1 | 1* |
| 1. Average Girls vs.<br>2. Slow Girls | | | | | | 1 | | 1 | | |

| | Sample IV | | Sample VI | | | Sample VII | | Sample VIII | | Sample IX |
|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 7 | 7 | 8 | 9 | 8 | 9 | 9 | 10 | 9 |
| District R | | | | | | | | | | |
| 1. Superior Boys vs.<br>2. Average Boys | 1 | 1 | | | | 1 | | | | 1 |
| 1. Superior Boys vs.<br>2. Slow Boys | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1. Average Boys vs.<br>2. Slow Boys | 1 | 1 | | 1 | | | | 1 | 1* | |
| 1. Superior Girls vs.<br>2. Average Girls | 1 | | 1* | | | 1 | | | | |
| 1. Superior Girls vs.<br>2. Slow Girls | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 |
| 1. Average Girls vs.<br>2. Slow Girls | 1 | 1 | 1 | | | | | 1 | | 1 |

* These differences approached statistical significance (t=1.65-1.95).

For District R pupils six Discrepancy Score differences reached or approached statistical significance, with five of these indicating higher (less favorable) Discrepancy Scores for the group of higher ability. Four of these six differences were between average and superior boys.

The results reflect a trend for superior boys in both districts to have somewhat less favorable Discrepancy Scores than boys of average or low ability. No other trends in the Discrepancy Score results seem to be sufficiently consistent to justify any conclusion.

Within-treatment comparisons on the Ideal Self score as summarized in Table 10-37 show a marked and highly consistent trend in both districts for pupils of higher ability to have significantly higher Ideal Self scores. In District A, 34 of the 60 within-treatment comparisons reached statistical significance. The same number of the within-treatment differences were significant in District R. In every instance these significant differences favored the group of higher ability. It may be concluded that in both ability grouped and random grouped classrooms there is a consistent relationship between the pupil's ability level and his Ideal Self score on the Index of Adjustment and Values.

## Self-Concept Changes Accompanying Ability Level Changes in District A

Pupils in District A, Sample IV who changed ability levels between the fifth and sixth grade administrations of the IAV and pupils in Sample VI who changed ability levels between the sixth and seventh grade administrations of the IAV were compared to determine whether changes in ability classification would be accompanied by changes in the self-concept variables. The results of these analyses may be found in Tables 10-12 and 10-18. The reader is reminded that the Sample VI data are generally more meaningful because sufficient cases were available to divide the subgroups by sex, and because the junior high school form of the IAV was administered at both sixth and seventh grade levels, yielding more comparable data than were obtained for Sample IV. From these analyses, it may be concluded that:

(1) Placement of a pupil in a lower ability classification appears to bring about lower scores in Concept of Self, Acceptance of Self, and Ideal Self.

(2) The IAV scores of girls assigned to a lower ability group appear to be more drastically affected than are the scores of boys under the same circumstances. This may reflect the tendency for girls to place a higher value on academic achievement and status.

(3) Pupils destined to be reassigned to a lower ability classification show lower scores in Concept of Self and Acceptance of Self prior to re-assignment. This lower initial score may indicate that misassignment of a pupil to a higher group than his ability would justify brings about a lowering of his Concept of Self and Acceptance of Self because of his inability to perform satisfactorally in his assigned group.

(4) Remaining in the same ability classification in an ability grouping system appears to be less threatening to the pupil's self-concept than reclassification to either a higher or lower section. Reclassification to a higher section may lead to a slight loss in the pupil's self-concept because he is probably near the bottom of the ability range in his new group. This self-concept loss may be compensated to some extent by the pupil's identification with the higher ability classification. The effect upon pupils assigned to a lower ability group, however, appears to be much greater. These pupils probably lose self-concept because of their initial mis-assignment and then their self-concept is further lowered, especially for girls, when they are

confronted with the stigma of identification with a lower ability group.

Whether the lower scores on the self-concept variables that occurred rather consistently in the ability grouped samples are considered "good" or "bad" depends on one's point of view. When considered in the light of psychological research in related areas such as aspiration level, however, the author must conclude that the apparent effects of ability grouping upon the self-concept variables that emerged from the Utah study are probably harmful to the development of at least some of the pupils who are educated under such a system of grouping.

## REFERENCES

Adler, A. *The practice and theory of individual psychology*. New York: Harcourt Brace, 1924.

Benjamins, J. Changes in performance in relation to influences upon self-conceptualization. *J. abnorm. soc. Psychol.*, 1950, 45, 473-480.

Bills, R. E., Vance, E. L., & McLean, O. S. An index of adjustment and values. *J. consult. Psychol.*, 1951, 15, 257-261.

Bills, R. E. *Index of Adjustment and Values Manual*. Auburn: Alabama Polytechnic Institute. Undated.

Bruck, M. A study of age difference and sex difference in the relationship between self-concept and grade-point average. *Dissertation Abstr.*, 1959, 19, 1946.

Conklin, A. N. *Failures of highly intelligent pupils*. Teachers College, Contributions to Education 792. New York: Bureau of Publications, Teachers College, Columbia University, 1940.

Drews, E. M. The effectiveness of homogeneous and heterogeneous ability grouping in ninth grade English classes with slow, average, and superior students. Unpublished manuscript. Michigan State University, 1962.

Freud, S. *The ego and the id*. London: Hogarth Press, 1950.

Fromm, E. Selfishness and self-love. *Psychiatry*, 1939, 2, 507-523.

Gough, H. G. *Manual for the California Psychological Inventory*. Palo Alto, California: Consulting Psychologists Press, 1957.

Horney, K. *The neurotic personality of our times*. New York: W. W. Norton, 1937.

Jersild, A. R. *In search of the self*. New York: Bureau of Publications, Teachers College, Columbia University, 1952.

Lecky, P. *Self-consistency: a theory of personality*. New York: Island Press, 1945.

Mann, M. What does ability grouping do to the self concept? *Childh. Educ.*, 1960, 36, 357-360.

Maslow, A. H. *Motivation and personality*. New York: Harper, 1954.

Perkins, H. V. Factors influencing change in children's self concepts. *Child Develpm.*, 1958, 29, 221-230.

Raimy, V. C. Self-reference in counseling interviews. *J. consult. Psychol.*, 1948, 12, 243-248.

Remmers, H. H., Shimberg, B., & Drucker, A. J. *Technical Supplement for the SRA Youth Inventory, Form A*. Chicago: Science Research Associates, 1953.

Remmers, H. H., & Bauernfeind, R. H. *Manual for the SRA Junior Inventory, Form S*. Chicago: Science Research Associates, 1955.

Rogers, C. R. *Client-centered therapy*. Boston: Houghton Mifflin, 1951.

Roth, R. M. Role of self-concept in achievement. *J. exp. Educ.*, 1959, 27, 265-281.

Simpson, R. E., & Martinson, R. A. Educational programs for gifted pupils. Sacramento, California: California State Department of Education, January, 1961.

Snygg, D., & Combs, A. W. *Individual behavior: a new frame of reference for psychology*. New York: Harper, 1949.

Thorpe, L. P., Clark, W. W., & Tiegs, E. W. *Manual, California Test of Personality, 1953 Revision*. Los Angeles: California Test Bureau, 1953.

Wylie, Ruth. *The self concept*. Lincoln, Nebraska: University of Nebraska Press, 1961.

## Chapter 11

# ASPIRATION LEVEL AND ACHIEVEMENT VALUES

by Walter R. Borg and George A. Jeffs[1]

## Introduction

Because of the apparent relationship between self-concept and aspiration level, this chapter has been placed immediately following discussion of the self-concept results so that the reader can consider these topics together. Probably because of the lack of aspiration level measures, excepting those of the laboratory type, no previous work has been published dealing directly with relationships between ability grouping and levels of aspiration. The lack of appropriate measures and previous research in this area necessitates that this phase of the Utah study be considered as exploratory. Measures dealing with classroom aspiration level, social aspiration level, and educational aspiration level were developed for this study. These are exploratory tools and considerable work would be required to build them into effective instruments and to establish their validity.

Pupils at the junior high school level were used as subjects in this research largely because younger pupils are unlikely to have realistic aspirations in some of the areas that were to be studied, and because treatment differences at the high school level were somewhat less clear cut.

Achievement Value (v-Achievement) may be said to be the degree to which a subject considers himself to be achievement oriented. This variable would appear to be rather closely related to both self-concept and level of aspiration and therefore data collected on this variable will also be discussed in this chapter.

Gardner (1940b) found that of the 51 adolescents investigated, the 10 with the highest aspiration level also rated highest on importance attached to intellectual achievement. One investigator (Miller, 1959) discovered a positive but low relationship between occupational aspiration and achievement values among senior high school boys. This same report indicated a low and non-significant correlation between an individual's achievement values and the achievement values of the members of his peer group. Haller and Miller (1961) discovered among high school boys a significant relationship between their Occupational Aspiration Scale and de Charms' v-Achievement measure. Pierce and Bowman (1960) found a significant difference in achievement values (using Strodtbeck's instrument) between tenth grade high and low achieving boys. A difference between these groups failed to reach the .05 level when de-Charms' instrument was used to detect achievement values. No significant differences in achievement value were noted between high and low achieving twelfth grade boys when either instrument was employed. In an investigation of family interaction, values, and achievement, McClelland, et. al. (1958) found that overachieving boys scored significantly higher than underachieving boys on needs for achievement and on achievement values.

---

[1] This chapter is based on Dr. Jeffs' Doctoral Dissertation carried out under Dr. Borg's supervision.

## WORKING HYPOTHESES

The hypotheses designed to be tested in this phase of the Utah study are concerned with between-treatment differences, i.e., differences between pupils of comparable ability in the ability grouped vs. random grouped districts; and within-treatment or level differences which are concerned with the differences between pupils at different ability levels within each district. The specific null hypotheses tested in this research were concerned with treatment and level differences for each of the five dependent variables.

As no previous research has been published that deals specifically with differences in aspiration level for pupils in ability grouped vs. random grouped classroom situations, the working hypotheses for this phase of the study are based primarily upon an appraisal of the psychological aspects of ability grouping and attempt to translate these psychological estimates into their possible effects upon the aspiration level of pupils in ability grouped and random grouped classes.

### Between-treatment Differences

With respect to superior pupils in District A, the added stress imposed by an accelerated program and the necessity to compete with pupils of comparable ability would seem to lead to more feelings of inadequacy and more frequent failures than would be experienced by superior pupils in heterogeneous classes. The research dealing with changes in level of aspiration following success and failure generally indicates that level of aspiration is raised following success and lowered following failure (Adams, 1939; Cohen, 1950; Festinger, 1942; Klugman, 1948; Pennington, 1940; Smith, 1949). This would suggest that, at least for laboratory type aspiration level measures of the sort used in more of the previous research, one might expect superior District A pupils to score somewhat lower than comparable pupils in District R.

The pressure described above would be expected to affect, perhaps to a lesser degree, the scores of superior pupils on other aspiration level measures. Certain aspects of classroom aspiration, for example, seem particularly subject to influence by the aforementioned difference introduced by ability grouped situation. Social aspiration might be expected to be reduced indirectly for superior pupils because of the reduced chances of the superior individual to obtain positions of leadership when placed in a classroom setting where he is called upon to compete with other superior pupils. Thus, with regard to educational and occupational aspiration and to a lesser degree, social aspiration, it appeared likely that aspiration level gains resulting from the pupil's identification with a superior group would be more than counteracted by the aforementioned variables that might be expected to lead to lowered aspiration levels for superior District A pupils. In view of the apparent relationship between self-concept and aspiration level, Drews (1962) finding that the self-concept of pupils became lower in homogeneous classes also supports this view. Thus, one of our working hypotheses was that, with respect to the aspiration level measures used in this research, District A superior pupils would score higher than comparable pupils in District R. Concerning value achievement, our working hypothesis was that the pressures of the accelerated system would tend to bring about higher achievement orientation for superior District A pupils.

With regard to the effects of treatment differences upon average pupils, our working hypothesis was that these differences would be far smaller for average pupils than for either superior or slow pupils, and as a result, no differences in aspiration level or value achievement would be found in comparisons between average pupils in District A and District R.

Many of the pressures that appear to operate on the superior pupils in the ability grouped classroom might be expected to operate in the opposite direction for slow pupils in the ability grouped classroom. For example, while competition becomes keener and failure more likely for superior pupils in ability grouped classrooms, the opposite might be expected for slow pupils who are working with a decelerated curriculum and competing in a situation where they would be expected to experience more frequent success than would be the case in the more heterogeneous classrooms of District R. Drews'research (1962) which showed much higher levels of participation for slow junior high school pupils in homogeneous English classes, than in heterogeneous classes, would lend support to the working hypothesis that the slow pupil would experience more recognition and success in his interactions in the ability grouped treatment.

While identification with the superior class might be expected to raise the individual's self regard and in turn lead to higher aspiration levels, identification with the slow group again might be expected to lead to the opposite result.

When all factors that might affect the aspiration level of slow pupils were weighed, the most acceptable working hypothesis appeared to be that ability grouping, by providing a more favorable climate for slow pupils than random grouping, would lead to higher laboratory and classroom aspiration levels, and higher value achievement. Similarly, the ability grouped classroom seems to provide a more favorable social climate for slow pupils. Thus, an increase in some aspects of social aspiration was also hypothesized.

With regard to educational and occupational aspiration, it would appear that the increased awareness of the individual's inferior achievement status as compared with his peers that is likely to occur in the ability grouped classroom would lead to a lowering of these aspects of aspiration level for slow pupils. The individual's identification of himself as a member of the slow group appears likely to lead to a more realistic or perhaps more pessimistic appraisal of his educational and occupational potentials than would be made by the slow pupil in the random grouped class where no such label is attached to him.

## Level Differences

As the predominant pressures in the ability grouped classrooms appear to lead to a lowering of classroom, laboratory, and social aspiration for superior pupils and an increase in these variables for slow pupils, it was hypothesized that no significant within-treatment differences would occur at these levels in District A. Conversely, it was hypothesized that within-treatment differences would occur in District R with superior pupils receiving the highest scores on these variables and slow pupils earning the lowest scores.

With regard to educational and occupational aspiration and value achievement, it was hypothesized that significant level differences would occur in both districts favoring the superior pupil as compared with the slow pupil. On all of the aspiration level measures, it was hypothesized that the average group would normally obtain scores between the superior and slow groups and would not be significantly different from either extreme.

## PREVIOUS RESEARCH PERTINENT TO ASPIRATION LEVEL AND ABILITY GROUPING

### Introduction

Although no studies were found in the leterature that were directly concerned with differences in aspiration level occurring in ability grouped and random group classroom situations, some work reported in related areas is sufficiently pertinent to have implications for the research reported herein. The three areas that appear to be most pertinent are studies dealing with: the influence of the group on aspiration level, the relationship between level of aspiration and achievement, and changes in aspiration level following success or failure. The review will therefore be confined to these areas.

### Group Influences on Level of Aspiration

Anderson and Brandt (1939) investigated the effects of the standards of one's own group upon his level of aspiration. One hundred fifth grade children were given a series of six cancellation tests spaced a half-week apart. The relative performance of each child was made available to him so he could see how he stood in relation to the group, but he could not identify the position of any other peer. Prior to each test the subjects were requested to write down privately the score they believed they could reach on the succeeding trial. Graphs showing relative standings were maintained throughout the experiment. Grouping the subjects according to performance quartiles gave lucid trends exposing the influence of knowledge of group standing. For the upper quartile (those scoring highest in the group) the level of aspiration was an average of 5.8 points below the performance level. For the second quartile, the level of aspiration was 1.9 points above the performance level. For the third quartile, the level of aspiration was 2.1 points above performance. For the fourth quartile, the level of aspiration was 13.6 points above the performance level. Thus, the trend was for those subjects finding themselves above the average of the group to have negative discrepancy scores, those finding themselves close to the average of the group tend to have slightly positive discrepancy scores, while those subjects who found themselves below the mean of the group tend to have very large positive discrepancy scores. This study found that the correlation between discrepancy score and position of performance with repsect to the group mean, was -.46. That is, the lower the performance relative to that of the group, the larger the discrepancy. Hilgard, Sait, and Margaret (1940) obtained essentially the same results in a study of college students.

Hertzman and Festinger (1940) explored level of aspiration among male college students. These subjects were given two experimental sessions. At the time of the second session, the subjects were told the average score and the average level of aspiration of a group of other people of their own scholastic standing. The scores assigned to this group were arranged so that the individual's performance was approximately equal to the performance of the group. The positive or negative sign of the

b

group goal discrepancy was reported as opposite to that of the subject. The results showed that the level of aspiration changed significantly from the first to the second session in the direction of conforming to the group level of aspiration. Rotter (1942a) concludes that an individual's expressed goal or aspiration level tends to regress toward the mean performance of the group to which he maintains association. This same investigator explains that a subject's knowledge of the performance of others tends to influence his difference score. A. B. Wilson (1959), studying high school boys, points to school milieu and peer group norms as having an influence upon educational and occupational aspirations. Haller and Butterworth (1960) report that the data gathered in their study on high school boys support the hypothesis that peers tend to influence each others' levels of educational aspiration. Another investigation among high school boys (Miller, 1959) disclosed a positive correlation of .33 between an individual's level of occupational aspiration and the aspiration levels of his peer group members.

Edwards and Wilson (1961) reported an interesting finding concerning peer group influence on level of aspiration:

> Peer groups of children from lower-class families are more tightly knit than are those from middle-class families. Hence, the peer group becomes the significant reference with respect to attitudinal development among children from the lower socio-economic sections of the population. (p. 112)

Showing the influence of school atmosphere on educational aspiration, these same investigators found that 93 per cent of the sons of professional people at high status schools desired education further than that offered by the high school and fewer than two-thirds of the sons of professional people at predominantly working-class schools wished to go to college.

One study (Zander and Medow, 1962), involving teams of three or five eleventh and twelfth grade boys, investigated group level of aspiration versus individual level of aspiration. These researchers offer the following important finding:

> Teams and solo persons alike raised their aspirations following a performance in which they exceeded a previously set level of aspiration. But teams more often lowered their aspiration following an unexpectedly poor performance than did solo persons. Thus, teams appeared to be less optimistic following a poor performance than were individuals. Teams appeared to be more concerned with the perceived probability of attaining a given level, when setting a level of aspiration, than were solo persons. (Zander and Medow, 1962, p. 22)

Chapman and Volkmann (1939), investigating college students, demonstrated that knowledge of what others did on a test (where the subject had not yet actually performed) will markedly influence the subjects' stated aspiration depending upon how they evaluate the other people whose scores they know. This study revealed that there is a tendency for subjects to aspire toward the middle of the distribution when they do not actually know what they themselves can do. Another investigation (Kaiser and Blake, 1955) revealed that an individual's level of aspiration may be changed by informing the subject that the group to which he claims membership has altered its goal desires. Rosenthal and Cofer (1948) showed that the reverse is also true. That is,an individual group member may alter the behavior of his fellow members as well as their goals.

Other investigations, using subjects of different age levels, offer results that tend to confirm the finding that the group to which an individual claims affiliation influences his level of aspiration (Dreyer, 1954; Festinger, 1942a, 1942b; Gilchrist, 1952; Gilinsky, 1949; Hanawalt, Hamilton and Morris, 1943; Klugman, 1948; Kostick, 1957; Margaret, 1942; Raven and Rietsema, 1957' Stotland, et. al., 1957; Veness, 1960; Zander, Natsoulas and Thomas, 1960).

## Level of Aspiration and Achievement

Achievement, as used in this review, shall refer to academic success taking form in grades, scores, honors and the like.

Sears (1940) conducted a much referred-to study involving children from ages 9 to 12 years. She divided these children into three groups. A "success" group was composed of children who had shown evidence of success in all of the academic school subjects including reading and arithmetic. A "failure" group was composed of children who had academic experience opposite that of the success group. A mixed group, called the "differential" group, was made up of children who had experienced success with reading but met failure with arithmetic. The tasks administered to these subjects were multiple-choice word-meaning items for reading, and problems in addition for arithmetic. Half of the subjects received arithmetic tasks during the first session and the other half received reading tasks. The reverse was true for the following session so that all the children were exposed to all of the tasks. Two sample trials were followed by 20 regular trials. After each trial the child was told his performance time, and then either spontaneously gave his next level of aspiration or was asked for it by the experimenter, who used the words: "Now what are you going to try for?" A success and failure condition was established for further tasks. The examiner for the success condition was especially supportive and encouraging. The environment for the failure condition was the opposite. It was shown that self-confident, successful children react to the level of aspiration situation in pretty much the same way, whereas unsuccessful children, lacking in confidence, may adopt behavior techniques that are inconsistent. Sears found that experimentally induced success brings the reactions of all subjects in regard to level of aspiration into a more homogeneous distribution than do the neutral conditions of stimulation. It was found that the children of the past failure group showed a higher goal discrepancy on the average than those of the past success group. More pronounced, however, was the wide variability among subjects of the failure group. The range of discrepancies among these subjects was from very high positive to negative scores. The variability among subjects of the success group was much less, with discrepancies almost entirely within a small positive range.

Sears (1940) relates that Juchnat grouped 500 subjects into categories of high, medium or low on the basis of their initial levels of aspiration. Juchnat then examined the grades of these subjects, and found that better students showed a high initial level of aspiration, average students a medium initial level, while the poor students set an initial level which was either at the top or bottom of the scale. Pierce and Bowman (1960), studying tenth and twelfth grade boys, found the high achievers to possess significantly higher educational goals than low achievers. A study involving 17 year old boys (Haller and Miller, 1961) disclosed a relationship between achievement (grade point average) and occupational aspiration of .50, while the correlation between achievement and educational aspiration reached .53. Anderson and Brandt (1939) found that fifth grade pupils in their study who were rated in the lowest levels of achievement consistently set goals considerably above their past achievement and pupils rated in the upper levels of achievement were consistent in setting goals somewhat below their preceding achievement. Kurtz and Swenson (1951), employed the term "plus achiever" to indicate a student whose achievement is well above expectation on the basis of ability rating and "minus achiever" to indicate a student whose achievement is definitely below expectation on the basis of ability rating. These researchers disclosed that plus achievers set high educational and vocational aims, and that they are likely to achieve their goals. Plus achievers appear to relate their school work to future goals, and to regard an education as important for more than its job value. Minus achievers have limited educational and vocational aims and do not look to the future. Plus achievers were shown to possess a stronger desire for an extended education and more concrete plans for attaining it. Plus achievers also appear more concerned about the possibility of not securing the amount of education they desire. Plus achievers have fairly clear vocational aims.

One investigation (Wilson, 1959) revealed that high school boys achieving highest grades possess higher educational aspirations than do students achieving lesser grades. Wilson also believes that socioeconomic status is an important factor in determining level of aspiration regardless of grade point average. Hilgard (1946) learned that children with a past history of success in reading and arithmetic usually established their goals realistically when examined in either of these areas. He found that children who consistently failed exhibited goal-setting behavior that was unpredictable. Another study (Hilgard, Sait, and Margaret, 1940) revealed that among college students, those succeeding in an experimental situation interpreted their success cautiously and made estimates that the next score would be less successful. Those college students who met failure in the experimental situation estimated that their scores would improve.

A study by Mitchell (1959) involving self-acceptance, achievement, and aspiration level of college females reported the following findings:

> (1) self rejectant overachievers consistently have the least discrepancy between level of aspiration and grade actually obtained, (2) self-rejectant overachievers are most likely to receive an estimation grade that satisfies or surpasses their aspiration, (3) self-rejectant overachievers always show the smallest discrepancy between prior grade and present level of aspiration, (4) self-acceptant underachievers show with consistency the greatest discrepancy between previous grade and present level of aspiration, and (5) self-acceptant underachievers always possess the largest discrepancy between expected grade and actual grade obtained. (pp. 102-103)

Barnett, Handelsman, Steward and Super (1952) found among junior and senior high school students a high relationship between level of occupational aspiration and academic achievement. This finding was supported by the work of Bradley (1943) and Stubbins (1950). An interesting discussion of the relationship between level of aspiration and achievement is given by Dorcus and Dunlap (1940):

> The dull child has to work up to his maximum all the time, whereas the bright child is not forced to exert himself to keep up to the class level. When special incentives are added, only the bright child can show

> improvement. Aspiration and achievement are mutually dependent. The level of aspiration is modified by success and failure, and experiences of success and failure are relative to the level of aspiration. The unpracticed golfer who hopes to shoot a round in 90 will experience the elation of success if he shoots 88, while the expert who usually shoots his round in the seventies will have had a terrible day if he requires 80 strokes. Hence, the effect of success or failure on an individual is dependent upon the aspiration level the individual sets for himself. The level of aspiration must be adjusted from time to time to provide a goal suitable to the development that takes place within the individual; otherwise no improvement would occur. If the goal is adjusted at a level beyond the reach of the individual's ability, the results are likely to be disastrous. (Dorcus and Dunlap, 1940, p. 289)

Worrell (1959) found that college students who held aspirations close to previous performance, who considered performance below their previous performance as satisfactory, and who believed they would not achieve greatly beyond their present achievement by applying greatest effort, tended to be successful in grade achievement. Pennington (1940) showed that failure among college students is frequently accompanied by a lowering of level of aspiration.

### Level of Aspiration Changes Following Success and Failure

It might be said that the goal discrepancy on level of aspiration measures indicates the general tendency of the subject to set his goals higher or lower than previous performance. The goal discrepancy score may not, however, indicate how the subject's goals vary from performance to performance when immediately preceding by success, or failure. Festinger (1942b) found that levels of aspiration are generally raised after success and lowered following failure. Several investigators have researched this same area and are generally in agreement with Festinger's findings (Adams, 1939; Cohen, 1950; Lysenck and Himmelweit, 1946; Klugman, 1948; Pennington, 1940; Smith, 1949). Zander and Medow (1962) found this relationship to be more pronounced when level of aspiration is established by the group rather than by an individual member of the group. The number of shifts either upward of downward and the times a subject failed to shift following both success and failure has also been experimentally investigated (Cohen, 1950; Rotter, 1945). Research on this same theme was conducted by Sears (1940) to determine the times that levels of aspiration are set equal to or lower than preceding performance (negative discrepancy), while Klugman (1948) investigated the same area but was concerned with positive discrepancy.

## PROCEDURE

### Subjects

### Introduction

The 554 subjects selected for the level of aspiration study were taken from the male pupils in Sample IV and VI. These students were enrolled in grades seven and nine during 1961-62, the school year in which the aspiration level data were collected.

One reason why the grade levels involved in the level of aspiration study were so selected was because little research concerning aspiration at these grade levels had been carried out. It will be recalled that at these grade levels in District A, ability grouping was employed for mathematics and science. The age level at which we find the subjects of this study is one at which educational, social, and occupational aspiration are only partially formulated. These are probably years in which these aspirations are being formulated and thus, if the treatment differences have an effect upon aspiration level, this effect should be reflected in these pupils. The samples had been exposed to the different grouping treatments for over three years at the time the aspiration level data were collected, thus allowing this period for any cumulative effects of the treatment differences to have their influence upon the pupils' formation of aspiration levels.

## District A Sample

District A pupils were assigned to the various mathematics and science sections, designated as 1, 2, 3, 4, for the 1961-62 academic year as a result of the scores received on the California Achievement Test battery. Table 11-1 gives an account of the number of subjects found at each level in District A. The first number of the level combination indicates the mathematics level at which the subject was enrolled. The second number of the level combination points to the science level to which the subject was assigned. A "1-1" indicates that the subject was enrolled in the superior section for both mathematics and science. A "1-2" combination indicates that the student was enrolled in a superior section for mathematics and an average section for science. A "2-1" combination indicates that the student was enrolled in an average section for mathematics and a superior section for science. A "2-2" combination means that the subject was enrolled in an average section for both mathematics and science. A "3" indicates the slow section while a "4" is indicative a special education section. The few students enrolled in a special education class in both districts were eliminated from all phases of the study as this research is concerned with pupils within the normal range of ability.

It may be noted from Table 11-1 that a limited number of subjects were found in the 2-3, 3-2, and 3-3 combinations for grade seven. These three levels were combined with 3-3 classification for the aspiration level study when t-tests revealed that there were no significant differences between the three groups on the aspiration level variables. This process increased the number in this group to 34. The 1-1 group at the seventh grade level included 31 subjects, the 2-2 level was composed of 69 subjects, and the adjusted 3-3 level was composed of 34 subjects.

The 3-3 section for grade nine included only 26 subjects. A t-test between sections 3-2 and 3-3 revealed no significant differences and thus these two levels were combined, giving 47 cases. It may be noted from Table 11-1 that the 1-1 group includes 32 subjects, the 2-2 group includes 72 subjects, and the adjusted 3-3 group consists of 47 subjects at the ninth grade level. All other subjects tested in District A were eliminated from the present study either because the N was too small for the particular subgroup or because they were in special education classes.

## District R Sample

It will be recalled that ability levels assigned to pupils in the District R sample were based upon cutoff scores on the STEP science and mathematics tests. These

Table 11-1

Number of Subjects at Each Level in District A
Grades Seven and Nine

| Level | Grade<br>Seven | Nine |
|---|---|---|
| 1-1 | 31 | 32 |
| 1-2 | 3 | 0 |
| 2-1 | 4 | 12 |
| 2-2 | 69 | 72 |
| 2-3 | 9 | 4 |
| 3-2 | 11 | 21 |
| 3-3 | 14 | 26 |
| 3-4 | 0 | 2 |
| 4-3 | 0 | 0 |
| 4-4 | 7 | 7 |
| Totals | 148 | 176 |

cutoffs were established by administering the STEP to pupils in District A at each ability level and using their scores to establish the District R cutoffs. Only pupils classified at the same ability level in both mathematics and science were used. There were 31 superior, 50 average and 25 slow pupils in the portion of the District R Sample IV used in this phase of the research. District R Sample VI pupils totaled 83, 41, and 39 for the three ability levels.

## Measures Employed

### Introduction

A total of five different measures of aspiration level plus a measure of value achievement were employed in this phase of the Utah study. As most aspiration level measures deal with what may be referred to as "laboratory aspiration," previously discussed in the review of the literature, and as these measures were inappropriate for exploring some of the aspiration level variables with which this project was concerned, it was necessary to develop some of the required measures and adapt others from previous research projects. Development and adaptation of these measures was carried out by Dr. Jeffs in a pilot study. The entire test battery was administered to 100 seventh grade boys and 140 ninth grade boys at a Utah junior high school similar to those found in District A and District R. Item analysis of the provisional battery led to making changes in the locally developed measures and further adaptations to some of the measures developed by others. The measures developed by the investigators for use in the Utah study, may be found in Appendix F.

### Classroom Aspiration

The classroom aspiration instrument is a self-appraisal rating that provides a measure of the subject's interpretation of his own skill at various performances together with the level of skill that he expects to reach. The areas covered by this instrument are sports, mathematics, science, reading, talking, and leadership, with a pair of questions devoted to each area. Most pupils complete this test in three minutes.

The classroom aspiration test is scored on the basis of the difference between the subject's estimate of his current status and his expectations in the same area. For example, should a subject consider himself "pretty good" at sports right now and expect to become "nearly the best in his class," his difference score (D-score) would be a plus 1. The D-score between each of six pairs of questions is calculated and these scores are summed for the final classroom aspiration score.

### Social Aspiration

The social aspiration instrument was designed to assess the expectations of the subjects in terms of social participation the year following high school graduation as well as five years hence. This questionnaire includes eight multiple choice items and is scored by assigning a different weight to each of the five alternatives offered as possible answers to each question. The first alternative is awarded five points, the second four points, the third three points, the fourth two points, and the fifth one point. Thus, it is possible to receive from one to five points for each question. Most pupils completed this test in seven minutes.

## Educational Aspiration

The educational aspiration instrument attempts to evaluate the degree of importance attached to different levels of academic progress for both the individual and his future family.

The higher the student's educational aims, the greater is the number of points he receives. The total educational aspiration score is the sum of scores obtained on each of the eight questions in the test.

## Occupational Aspiration

The instrument employed in the study for the purpose of investigating occupational aspiration was the Occupational Aspiration Scale (OAS), developed by Haller (1958). The OAS consists of eight forced-choice questions regarding job preference. This instrument is designed to estimate the occupational prestige level sought by the subject while minimizing the non-prestige aspects of his particular occupational choice. The examinee is required to select from 10 ranked occupational alternatives, one for each question.

Scoring is accomplished by assigning a weight to the alternative selected by the subject.

## Value Achievement

The Value Achievement test is the only measuring device employed in this study that does not directly relate to the level of aspiration. The purpose for using this instrument in the Utah study was to determine the effects of grouping on one's evaluation of achievement. The investigators are indebted to Richard deCharms and Irvin W. Miller for the use of the v Achievement questionnaire. The v-Achievement tool consists of nine questions designed to assess the degree to which an individual values achievement or believes himself to be achievement oriented. This instrument is so conducted that it indicates the student's opinion regarding his achievement motivation. Eight of the nine questionnaire items come from Murray's original study. Dr. deCharms and his associates (1955) found that the nine items that compromise the v-Achievement test have item validities of at least .30. The last item of this measure was rephrased to better insure intelligibility among seventh and ninth grade students in the Utah study. This item as presented by deCharms read, "Only ambition will bring a man's mind into full activity." The item as reworded by the investigators reads, "I think that people should be ambitious because only ambition makes one's mind active." The nine items are dispersed among 11 other items in an attempt to somewhat conceal the fact that this was a measure to determine the degree to which achievement is valued. The time involved in the administration of this questionnaire exclusive of instructions is approximately six minutes.

A score for the v-Achievement instrument is obtained by summing the weighted responses for each of the nine items fundamental to the questionnaire. The range of response categories is from strongly disagree (SD) to strongly agree (SA) and the weights assigned are from zero to four. Thus, high scores are indicative of greater achievement orientation. The total v-Achievement score may range from zero to 36.

### Laboratory Aspiration

The investigators are indebted to Dr. Raymond B. Cattell for the use of the Coding Test. The Coding Test is much like many other so-called laboratory aspiration tests, and may be classified as cognitive, nonverbal. The aims of this test are to get (a) a measure of the general excess of aspiration over achievement, (b) a measure of the immediate responsiveness of aspiration to changes in achievement and (c) a measure of aspiration for improvement, or the constant deficit between immediate past performance and immediately succeeding goal. The last aim is designed for children only between ages 10 and 17. The present study was concerned with measuring this aim. The task involved completing a coding test that requires substituting symbols for numbers. Six minutes were required to administer this test after directions had been read.

A score for the coding test is obtained by taking the sum of the aspiration "estimates" on the first two pages and subtracting this from the sum of the "estimates" on the last two pages. The estimate on page three is not directly used. It should be noted that a constant of 50 was added to each raw score as a means of eliminating negative scores.

## RESULTS

### Introduction

A single classification analysis of variance was carried out to test the hypotheses for mean differences on each of the dependent variables. F-tests were obtained from the analysis of variance and, when significant, were followed by t-tests to determine the significance of differences between comparable groups. It will be recalled that data on the dependent variables were collected for Sample IV and Sample VI pupils during their seventh and ninth grade years respectively. These data provide a cumulative estimate of the possible effects of the two grouping systems upon the dependent variables as both samples were tested during the final year of the project. Some differences in the results for District A pupils in Sample IV and VI might be expected, however, in view of the fact that the majority of Sample IV pupils' experience with ability grouping took place in the elementary school while the experience of Sample VI pupils was predominately at the junior high school level where ability grouping was employed in mathematics and science classes only.

### Between Treatment Differences (Sample IV)

Both between-treatment and level differences for Sample IV on the six dependent variables are summarized in Table 11-2. In examining this table it will be noted that only three differences significant beyond the .05 level are present among the 18 between-treatment comparisons. There are no significant between-treatment differences for Sample IV in classroom aspiration level, educational aspiration level, or value achievement.

In general, differences between comparable District A and District R groups on the aspiration level measures were very small. All three of the significant differences occurred between average groups in the two districts with District R average pupils obtaining higher average scores on social aspiration and occupational aspiration and District A average pupils obtaining a higher mean score on the laboratory aspiration

## Table 11-2

Aspiration Level Means for Sample IV, During Their 7th Grade Year

| | Between-Treatments | | | Level | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Ability Grouped | | | Randomly Grouped | | | |
| Measure | Accelerated | Average | Developmental | | | | | | | Grand Mean |
| | 1 vs 4 | 2 vs 5 | 3 vs 6 | 1 vs 2 | 1 vs 3 | 2 vs 3 | 4 vs 5 | 4 vs 6 | 5 vs 6 | |
| Classroom aspiration[a] | 13.741<br>14.387 | 15.275<br>15.300 | 14.764<br>15.600 | 13.741**<br>15.275 | 13.741<br>14.764 | 15.275<br>14.764 | 14.387<br>15.300 | 14.387<br>15.600 | 15.300<br>15.600 | 14.92 |
| Social aspiration | 17.000<br>16.451 | 14.826<br>16.720* | 14.264<br>13.480 | 17.000**<br>14.826 | 17.000**<br>14.264 | 14.826<br>14.264 | 16.451<br>16.720 | 16.451**<br>13.480 | 16.720*<br>13.480 | 15.49 |
| Educational aspiration | 33.645<br>34.774 | 31.898<br>32.760 | 31.705<br>29.720 | 33.645<br>31.898 | 33.645<br>31.705 | 31.898<br>31.705 | 34.774<br>32.760 | 34.774**<br>29.720 | 32.760**<br>29.720 | 32.42 |
| Occupational aspiration | 49.161<br>53.032 | 42.028<br>46.380* | 41.558<br>37.600 | 49.161**<br>42.028 | 49.161**<br>41.558 | 42.082<br>41.558 | 53.032*<br>46.380 | 53.032**<br>37.600 | 46.380**<br>37.600 | 44.75 |
| Value achievement | 22.451<br>23.483 | 21.159<br>21.740 | 20.000<br>20.720 | 22.451<br>21.159 | 22.451<br>20.000 | 21.159<br>20.000 | 23.483<br>21.740 | 23.483<br>20.720 | 21.740<br>20.720 | 21.53 |
| Laboratory aspiration | 62.580<br>60.258 | 63.202*<br>58.740 | 60.382<br>55.880 | 62.580<br>63.202 | 62.580<br>60.382 | 63.202<br>60.382 | 60.258<br>58.740 | 60.258<br>55.880 | 58.740<br>55.880 | 60.65 |

1=ability grouped, accelerated
2=ability grouped, average
3=ability grouped, developmental
4=randomly grouped, accelerated
5=randomly grouped, average
6=randomly grouped, developmental
**1 percent level of confidence
*5 percent level of confidence

[a] On this measure only a low score indicates higher aspiration.

Table 11-3

Directional Trends in Aspiration Level Measures for Sample IV

| | | | | Level | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Between-Treatment | | | Weber | | | Ogden | | | |
| Measure | Accelerated | Average | Developmental | 1 vs 2 | 1 vs 3 | 2 vs 3 | 4 vs 5 | 4 vs 6 | 5 vs 6 | Sign Test |
| Classroom aspiration | x | x | x | + | + | | + | + | + | 1.63 |
| Social aspiration | x | o | x | + | + | + | | + | + | 1.63 |
| Educational aspiration | o | o | x | + | + | + | + | + | + | 2.45* |
| Occupational aspiration | o | o | x | + | + | + | + | + | + | 2.45* |
| Value achievement | o | o | o | + | + | + | + | + | + | 2.45* |
| Laboratory aspiration | x | x | x | | + | + | + | + | + | 1.63 |
| Sign test | .000 | .816 | 1.63 | 1.63 | 2.45* | 1.63 | 1.63 | 2.45* | 2.45* | |

**1 percent level of confidence
*5 percent level of confidence
x=score favors ability grouped students
o=score favors randomly grouped students
+=indicates that higher academic level score is greater than comparison score.

measure. A sign test on the direction of between-treatment differences for pupils at each of the three ability levels was carried out and revealed no significant trends (see Table 11-3).

### Between-Treatment Differences (Sample VI)

Table 11-4 summarizes both treatment and level differences for Sample VI on the six dependent variables. It will be noted that the scores of comparable pupils in the two districts are very similar on these variables. Only two significant treatment differences emerged: District R slow pupils obtained more favorable scores on both classroom aspiration and educational aspiration than did slow pupils in District A. These two differences were small although reaching significance at the .05 level.

As 18 treatment comparisons were made for each sample, one difference significant at the .05 level would be expected by chance. Only three significant differences were obtained for Sample IV and two for Sample VI and there was no significant difference that occurred in both samples Thus, it appears that the most logical conclusion is that ability grouping and random grouping as used in the Utah study have no differential effect upon the measures of aspiration level and value achievement employed. It will be recalled that the measures used in this phase of the research were, for the most part, experimental and at best could be expected to provide only rough indications. One would expect, however, that if ability grouping or random grouping caused any marked differences in the variables concerned, these differences would have been found even with such crude measures as were available for use in this research.

### Level Differences

Comparisons between pupils at different ability levels also failed to reveal any clearcut results. No significant level differences occurred within either district on the measures of value achievement and laboratory aspiration level. In classroom aspiration level, three significant differences were found in the two replications, all between pupils in District A, and all favoring the pupils of higher ability. No significant classroom aspiration level differences occurred among the pupils of different ability in District R.

On social aspiration, four significant differences were found for Sample IV, two of these in District A and two in District R. In all four of these differences, the group of higher ability received the higher score. For Sample VI, however, none of the differences in social aspiration were statistically significant.

Educational aspiration results revealed two significant differences in Sample IV, both occurring in District R. Four of the six educational aspiration level differences were statistically significant for Sample VI with all three of the District A comparisons reaching significance at the .01 level. Again, the differences in both districts favored the pupils of higher ability level but in general, these differences were not consistent between Sample IV and Sample VI.

Occupational aspiration level differences between pupils of different ability levels were generally significant for both districts. In Sample IV, five of the six comparisons made reached statistical significance while four of the six differences in Sample VI reached significance. These results would indicate that occupational aspiration level is closely tied to ability but that the tie is no stronger in the ability grouping situation than it is in the random grouping situation.

## Table 11-4

Aspiration Level Means for Sample VI, During Their 9th Grade Year

| | | | | Level | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Between-Treatment | | | Ability Grouped | | | Randomly Grouped | | | |
| Measure | Accelerated | Average | Developmental | | | | | | | Grand |
| | 1 vs 4 | 2 vs 5 | 3 vs 6 | 1 vs 2 | 1 vs 3 | 2 vs 3 | 4 vs 5 | 4 vs 6 | 5 vs 6 | Mean |
| Classroom aspiration[a] | 13.343 | 14.166 | 15.510 | 13.343 | 13.343** | 14.166** | 13.530 | 13.530 | 14.170 | 14.14 |
| | 13.530 | 14.170 | 14.384* | 14.166 | 15.510 | 15.510 | 14.170 | 14.384 | 14.384 | |
| Social aspiration | 15.687 | 15.791 | 15.170 | 15.687 | 15.687 | 15.791 | 15.036 | 15.036 | 15.390 | 15.31 |
| | 15.036 | 15.390 | 14.820 | 15.791 | 15.170 | 15.170 | 15.390 | 14.820 | 14.820 | |
| Educational aspiration | 36.000 | 33.097 | 30.255 | 36.000** | 36.000** | 33.097** | 35,385 | 35.385** | 33.878 | 33.58 |
| | 35.385 | 33.878 | 32.358* | 33.097 | 30.255 | 30.255 | 33.878 | 32.358 | 32.358 | |
| Occupational aspiration | 50.812 | 45.152 | 41.000 | 50.812* | 50.812** | 45.152** | 51.277 | 51.277** | 47.731 | 47.04 |
| | 51.277 | 47.731 | 45.025 | 45.152 | 41.000 | 41.000 | 47.731 | 45.025 | 45.025 | |
| Value achievement | 23.375 | 23.083 | 22.404 | 23.375 | 23.375 | 23.083 | 22.120 | 22.120 | 21.804 | 22.25 |
| | 22.120 | 21.804 | 22.589 | 23.083 | 22.404 | 22.404 | 21.804 | 22.589 | 22.589 | |
| Laboratory aspiration | 61.250 | 64.069 | 58.957 | 61.250 | 61.250 | 64.069 | 60.686 | 60.686 | 60.853 | 61.29 |
| | 60.686 | 60.853 | 60.794 | 64.069 | 58.957 | 58.957 | 60.853 | 60.794 | 60.794 | |

1=ability grouped, accelerated
2=ability grouped, average
3= ability grouped, developmental
4=randomly grouped, accelerated
5=randomly grouped, average
6=randomly grouped, developmental
**1 percent level of confidence
*5 percent level of confidence
[a] On this measure only, a low score indicates higher aspiration.

## Table 11-5

## Directional Trends in Aspiration Level Measures for Sample VI

| | | | | Level | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | District | | | Weber | | | Ogden | | | |
| Measure | Accelerated | Average | Developmental | 1 vs 2 | 1 vs 3 | 2 vs 3 | 4 vs 5 | 4 vs 6 | 5 vs 6 | Sign Test |
| Classroom aspiration | x | x | o | + | + | + | + | + | + | 2.45* |
| Social aspiration | x | x | x | | + | + | | + | + | .816 |
| Educational aspiration | x | o | o | + | + | + | + | + | + | 2.45* |
| Occupational aspiration | o | o | o | + | + | + | + | + | + | 2.45* |
| Value achievement | x | x | o | + | + | + | + | | | .816 |
| Laboratory aspiration | x | x | o | | + | + | | | + | |
| Sign test | 1.63 | .816 | 1.63 | .816 | 2.45* | 2.45* | .816 | .816 | 1.63 | |

**1 percent level of confidence
*5 percent level of confidence
x=score favors ability grouped students
o=score favors randomly grouped students
+=indicates that higher academic level score is greater than comparison score.

## SUMMARY AND CONCLUSIONS

In summary, it will be noted that in every case where significant differences existed between pupils of different ability levels, the higher ability pupils received the higher score. The sign test was used to test the significance of this trend for each dependent variable in Sample IV and VI and it will be noted on Tables 11-3 and 11-5 that a number of these sign tests reached statistical significance. Significant trends were found in educational and occupational aspiration for both samples.

No characteristic pattern of level differences emerged that would tend to differentiate District A from District R. The only pattern emerging, in fact, was the general tendency for more able pupils to have higher aspiration levels and this tendency was reflected about equally in the two treatments. When the data on level and treatment differences are studied as a whole, they seem to support generally the aforementioned conclusion that the differences between ability grouping and random grouping occurring in this research appear to have had no consistent effect upon the measures of aspiration level or value achievement used.

## REFERENCES

Adams, D. K. Age, race and responsiveness of levels of aspiration to success and failure. Psychol. Bull., 1939, 36, 573.

Anderson, H. H., & Brandt, H. F. Study of motivation involving self-announced goals of fifth grade children and the concept of level of aspiration. J. soc. Psychol., 1939, 10, 209-232.

Barnett, G. J., Handelsman, I., Stewart, L. H., & Super, D. E. The occupational level scale as a measure of drive. Psychol. Mongr., 1952, 66 (10), 342.

Bradley, W. A., Jr. Correlates of vocational preferences. Genet. Psychol. Mongr., 1943, 28, 99-159.

Chapman, D. W., & Volkmann, J. A social determinant of the level of aspiration. J. abnorm. soc. Psychol., 1939, 34, 225-238.

Cohen, Louis D. Patterns of response in level of aspiration tasks. Educ. Psychol. Measmt., 1950, 10, 664-684.

de Charms, R., Morrison, W., Reitman, W., & MacClelland, D. C. The achievement motive. In David C. MacClelland (Ed.) Appleton-Century-Crofts, Inc., New York, 1955, 415-422.

Drews, E. M. The effectiveness of homogeneous and heterogeneous ability grouping in ninth grade English classes with slow, average, and superior students. Unpublished manuscript. Michigan State University, 1962.

Dreyer, A. S. Aspiration behavior as influenced by expectation and group comparison. Human Relat., 1954, 7, 175-190.

Edwards, T. B., & Wilson, A. B. A study of some social and psychological factors influencing educational achievement. Final report to the United States Office of Education, Department of Health, Education and Welfare, 1961. Project SAE 7787. (Mimeographed)

Eysenck, H. J., & Himmelweit, H. T. An experimental study of the reactions of neurotics to experiences of success and failure. J. gen. Psychol., 1946, 35, 59-75.

Festinger, L. A theoretical interpretation of shifts in level of aspiration. Psychol. Rev., 1942, 49, 235-250. (a)

Festinger, L. Wish, expectation, and group standards as factors influencing level of aspiration. J. abnorm. soc. Psychol., 1942, 37, 184-200. (b)

Gardner, J. W. The relation of certain personality variables to level of aspiration. J. of Psychol., 1940, 9, 191-206.

Gilchrist, J. C. The formation of social groups under conditions of success and failure. J. abnorm. soc. Psychol., 1952, 47, 174-187.

Haller, A. O., & Butterworth, C. E. Peer influences on levels of occupational and educational aspirations. Soc. Forces, 1960, 38, 289-295.

Haller, A. O., & Miller, I. W. The occupational aspiration scale: theory, structure, and correlates of an instrument designed to measure differential levels of occupational aspiration. Michigan State University, (Mimiographed), 1961.

Hanawalt, N. G., Hamilton, Carol E., & Morris, Louise M. Level of aspiration in college leaders and nonleaders. J. abnorm. soc. Psychol., 1943, 38, 545-548.

Hertzman, M., & Festinger, L. Shifts in explicit goals in a level of aspiration experiment. J. exp. Psychol., 1940, 27, 439-452.

Hilgard, E. R. Aspirations after learning. Childh. Ed., 1946, 23, 115-118.

Hilgard, E. R., Sait, E. M., & Margaret, G. Ann. Level of aspiration as affected by relative standing in an experimental social group. J. exp. Psychol., 1940, 27, 411-421.

Kaiser, R. L., & Blake, R. R. Aspiration and performance in a simulated group atmosphere. J. soc. Psychol., 1955, 42, 193-202.

Klugman, S. F. Emotional stability and level of aspiration. J. gen. Psychol., 1948, 38, 101-118.

Kostick, M. M. An experiment in group decision. J. teacher Educ., 1957, 8, 67-72.

Kurtz, J. J., & Swenson, Esther J. Factors related to over-achievement and underachievement in school. Sch. R., 1951, 59, 472-480.

Margaret, Ann. Influence of active participation in a social group upon expressed level of aspiration. Psychol. Bull., 1942, 39, 477.

MacClelland, D. C., Baldwin, A. L., Bronfenbrenner, U., & Strodtbeck, F. L. Talent and society. New York: D. Van Nostrand Company, Inc., 1958.

Miller, I. W. Occupational aspiration, v-achievement, and peer group membership. Department of Sociology and Anthropology, Michigan State University, 1959. (Mimeographed)

Mitchell, J. V., Jr. Goal-setting behavior as a function of self-acceptance, over- and underachievement, and related personality varibles. J. educ. Psychol., 1959, 50, 93-104.

Pennington, L. A. Shifts in aspiration level after success and failure in the college classroom. J. gen. Psychol., 1940, 23, 305-313.

Pierce, J. V., & Bowman, P. H. Motivation patterns of superior high school students. U.S. Department of Health, Education, and Welfare, OE-35016. The gifted student. Monograph No. 2, 1960.

Raven, B. H., & Reitsema, J. The effects of varied clarity of group goal and group path upon the individual and his relations to his group. Hum. Relat., 1957 10, 29-45.

Rosenthal, D., & Cofer, C. N. The effect on group performance of an indifferent and neglectful attitude shown by one group member. J. exp. Psychol., 1948, 38, 568-577.

Rotter, J. B. Level of aspiration as a method of studying personality: I. A critical review of methodology. Psychol. R., 1942, 49, 463-474.

Rotter, J. B. Level of aspiration as a method of studying personality: IV. The analysis of patterns of response. J. soc. Psychol., 1945, 21, 159-177.

Sears, Pauline S. Levels of aspiration in academically successful and unsuccessful children. J. abnorm. soc. Psychol., 1940, 35, 498-536.

Smith, C. H. Influence of athletic success and failure on the level of aspiration. Res. Quart. Amer. Ass. Hlth., 1949, 20, 196-208.

Stotland, E., Thorley, S., Thomas, E., Cohen, A. R., & Zander, A. The effects of group expectations and self-esteem upon self-evaluation. J. abnorm. soc. Psychol., 1957, 54, 55-63.

Stubbins, J. The relationship between level of vocational aspiration and certain personal data. *Gen. Psychol. Monogr.*, 1950, 41, 327-409.

Veness, T. Goal-setting behavior, anxiety and school streaming. *Brit. J. educ. Psychol.*, 1960, 30, 22-29.

Wilson, A. B. Residential segregation of social classes and aspiration of high school boys. *Amer. soc. R.*, 1959, 24, 836-845.

Wilson, J. A. R. Achievement, intelligence, age and promotion characteristics of students scoring at or below the tenth percentile on the California Test of Personality. *J. educ. Res.*, 1959, 52, 283-292.

Worell, L. Level of aspiration and academic success. *J. educ. Psychol.*, 1959, 50, 47-54.

Zander, A., & Medow, H. Individual and group levels of aspiration. Unpublished report to appear in *Hum. Relat.*, 1962.

Zander, A., Natsoulas, T., & Thomas, E. J. Personal goals and the group's goals for the member. *Hum. Relat.*, 1960, 13, 333-344.

# Chapter 12

# PERSONALITY

## INTRODUCTION

Although many of the arguments heard among educators on the merits and demerits of ability grouping imply that this practice can have an effect upon the personality development of pupils involved, very little direct research evidence comparing the personality characteristics of pupils in ability and random grouping systems has been published up to the time of this writing. Limited bits of indirect evidence such as the research carried out by Luchins and Luchins (1948) and the work of Mathias (1959) at least suggest the possibility that ability grouping and random grouping can lead to differences in personality development. Subjective appraisal of the differences that appear to be present in ability grouped and random grouped classes along with the consideration of the limited evidence on the relative merits of these systems suggest possible personality differences that might occur as a result of differences in grouping treatments.

It might seem a simple matter to identify a test measuring the qualities that we hypothesize are affected by the different grouping treatments, administer the test to pupils in the two treatments and determine whether differences between the scores were statistically significant. This is essentially the approach used in this project, but the reader should be aware of the many limitations attending this approach in the area of personality measurement. First, the greatest differences between the grouping systems studied in this research were at the elementary school level where the self-contained classroom system resulted in placing District A pupils in a single ability level classification for the entire school day. Thus, if differences were expected to occur as a result of the grouping treatments, these differences might be expected to show up most clearly at the elementary school level. Unfortunately, however, at this level very few personality measures are available, and those available can be regarded at best as rather crude instruments. At the junior high school level, a greater number of personality measures are available but it will be recalled that, as District A employed ability grouping only in mathematics and science at this level, the treatment differences were much less marked than at the elementary level and consequently would be expected to bring about smaller differences in personality.

Another practical problem arising in the study of personality differences between pupils in the two grouping treatments is the question of how much change in personality might we expect to take place as a result of changing a fairly small portion of the individual's total experience. Let us suppose, for example, that one grouping treatment did tend more frequently than the other to develop inferiority feelings. It would seem a strong possibility that the many other experiences in the pupil's day that may structure personality and that cannot adequately be controlled could mask any such personality differences that resulted from the grouping treatment.

A related question that must be faced is whether or not personality changes that might be brought about by the two grouping treatments could be expected to become apparent in a period of one or two years? The lack of any basis for answering this question, of course, leaves us with the possibility that the generally negative results obtained could reflect an inadequate exposure of the subjects to the treatment differences.

Another factor to be considered is the generally low level of our knowledge concerning the nature of personality and its measurement. Evidence of validity is rather scant for most of the published personality inventories, and even when validity data are given these data usually are based upon correlations with clinical judgment or with other criteria that themselves are of questionable validity.

Some psychologists, though condemning the personality inventory, place great confidence in projective tests as a source of personality information. This project employed both group administered personality inventories and individual and group administered projective measures. The unbiased researcher, however, must conclude that neither the personality inventory nor the projective measure at its current state of development provides very solid evidence. Thus, the personality findings presented in this and the following chapter should be considered at best tentative.

## WORKING HYPOTHESES

### California Test of Personality

Previous research evidence provided little information of help in establishing working hypotheses for the sub-tests of the California Test of Personality and these hypotheses are therefore quite tentative. For the first of these sub-tests, Feeling of Belonging, it was hypothesized that pupils at both the superior and slow ability levels would obtain more favorable scores in ability grouped classrooms. It would appear that within the school situation, these pupils would experience a greater feeling of belonging in the less heterogeneous District A groups. Slow pupils, especially, would seem more likely to be secure in a group of other slow pupils. A gradual deterioration of this characteristic for pupils in slow groups might occur, however, outside of the classroom. Ability grouping may cause both bright and slow children to sense that they are "different." For the slow child this difference carries with it the label of inferiority that is implied by his membership in the low ability class. Thus, the very act of separating pupils into groups on the basis of ability might make them actively aware of differences which could in turn lower their "feeling of belonging."

Both Withdrawing Tendencies and Anti-social Tendencies might be reactions to frustration. Evidence obtained in Drews' study (1962) that showed much higher levels of class participation for slow, ability grouped students would suggest that such students would be less likely to display these tendencies in ability grouped classes. Studies by Jackson (1943) and Marsh (1955) also reported fewer anti-social behavior patterns in the ability grouped situation, but presented no quantitative data to support their findings. On the other hand, however, the labeling of children in low ability classes seems to present a situation where both withdrawal and anti-social reactions could be expected. In light of the rather limited evidence, however, the working hypothesis of this research was that low ability pupils would display significantly more favorable scores on these measures in the ability group situation.

Mathias' work, (1959) which is discussed in the next section of this chapter, would also suggest that less aggression would occur in ability grouped classes.

Our working hypothesis with regard to average pupils on these three variables and to superior pupils on the latter two was that no significant between-treatment differences would occur.

## Inventory of Factors GAMIN

It was hypothesized that scores on Factor A, ascendancy in social situations, would be higher for superior pupils in District R because of the greater opportunities for these pupils to obtain positions of social leadership in random grouped classrooms. Slow pupils, on the other hand, would appear to have a greater opportunity for social leadership in the ability grouped classroom. Thus, it was hypothesized that slow pupils in District A would obtain higher Factor A scores than comparable pupils in District R. As ability grouping would seem to discourage ascendant behavior in superior pupils and encourage it in slow pupils, it was hypothesized that no significant differences between District A pupils at different ability levels would be found on this factor. Conversely, in District R it was hypothesized that differences between pupils of different ability levels would reflect a positive relationship between ability and ascendancy in social situations.

The other score obtained on the Inventory of Factors GAMIN was Factor I which purports to measure self-confidence or lack of inferiority feelings. It was hypothesized that District A slow pupils would receive less favorable scores on this factor than comparable pupils in District R. It was also hypothesized that superior District A pupils would receive less favorable scores than superior District R pupils although this difference might be expected to be less pronounced than the difference between the slow groups. For the slow child, many educators have suggested that the labeling that accompanies ability grouping would tend to lower self-confidence and develop inferiority feelings. Some rather limited research evidence supports this contention (Luchins and Luchins, 1948). The superior pupil might also be expected to have greater self-confidence in the random grouped class because of the smaller number of pupils in the class able to compete with him academically and the greater chances he has of obtaining a position of social recognition within the peer group.

## Objective Analytic Personality Factors

Two factors from the Objective Analytic Personality Test Batteries were measured in the Utah study. Cattell (1955) interprets Factor UI-16, Competent Assertiveness, as follows:

> This title, like those of the other factors, is a compromise between sheer description and tentative interpretations. The individual high in this factor shows performances which suggest a competitive assertiveness. He does a whole range of things well--exploration, fluent expression, showing taste in reading, the C.M.S., etc.--yet the dimension is not general intelligence. It could well be that we have here the main factor in the "need for achievement" phenomena studied by McClelland. Whenever fast, determined, independent, individualistic, poised action is involved the factor should be predictive. It has been found, in fact, to be positively correlated with ratings

> of total effectiveness, by superiors, of pilots in training, and might be still more correlated with salesmanship and success in situations requiring social aggressiveness. (p. 52)

It was hypothesized that superior pupils in District R would obtain higher scores on this factor than superior pupils in District A. The random grouping situation would seem more likely to develop the self perceptions underlying this factor in superior pupils than would the ability grouping situation.

Regarding within-treatment differences, it would appear likely that this factor would be positively related to ability level in both districts with greater differences occurring between ability levels in District R than in District A. No working hypothesis regarding sex differences was established for this variable.

Cattell's interpretation of Factor UI-24, Anxiety to Achieve, is as follows:

> The fact that this dimension corresponds to the clinical concept of anxiety may not be self evident from its internal loadings alone though high irritability, emotionality of comment and uncertainty about untried performances are consistent with all three of the oblique questionnaire factors of anxiety C-, O- and Q4.
>
> It will be observed that the anxiety measure is also loaded with verbal ability, good school grades, (intelligence being constant) and good ethical choices. Accordingly, it can be interpreted as a general anxiety (not a neurotic anxiety alone) connected with acceptance of cultural pressure and maintenance of a high degree of long-circuiting of impulses in the interests of achievement and cultural values. It is the degree of internal stress produced in the person either by his own goals or by a threatening situation. It would be of value especially in clinical and psychosomatic research, and in prediction of scholastic achievement. It has a slight positive correlation even with flying school success. (Cattell, 1955, pp. 54-5)

## California Psychological Inventory

Working hypotheses for the California Psychological Inventory were developed for each of the four classes of measures as shown in Table 12-1.

Class I includes six measures in the broad area of poise, ascendancy and self assurance. One of these areas, Self Acceptance, was reported separately along with the other self-concept data in Chapter 10. For the Class I variables, it was hypothesized that low ability pupils would score higher on these dimensions in the ability grouped classroom, superior pupils would score higher in the random grouped classrooms and no differences would be found for average pupils. The ability grouping situation seems to provide the low ability pupil considerably more opportunity for developing some of the Class I characteristics. Dominance, for example, seems to have a much greater chance of emerging in a situation where the slow pupil is participating with peers of similar ability. Differences in grouping treatments, might also be expected to have some influence on the slow pupil's Capacity for Status and Sociability and should contribute directly to his Social Presence. This is supported by Drews' research (1962) which indicated that the slow pupil has more opportunity to experience successful social interaction in the ability grouped situation. The final Class I variable, Sense of Well Being seems less clear on its possible effects upon the slow pupil.

Ability grouping, on the other hand, appears to create an unfavorable climate for superior pupils to develop the Class I personality traits. When removed from their favored position in the random grouped classroom, it would appear that the superior pupil's opportunities to develop Dominance, Social Presence, and Sense of Well Being, would be reduced. Sociability might similarly be reduced for the superior pupil in the ability grouped situation as it would seem likely that he would be less secure, less able to dominate the social situation and less able to gain social experience in the more homogeneous class.

No working hypotheses were developed for the six measures of socialization, maturity and responsibility which make up the Class II variables on the CPI. These scales were administered largely because of the advantages of using the entire test as opposed to selecting specific scales as was done with the California Test of Personality. Data for the Class II measures were analyzed, however, and will be presented.

Three of the scales on the CPI make up Class III--measures of achievement potential and intellectual efficiency. For two of these measures, Achievement via conformance (Ac) and Intellectual efficiency (Ie) it was hypothesized that superior pupils in ability grouped classrooms would receive higher scores. Ability grouping with acceleration where pupils are exposed to a uniform program at a rate close to their capacity would seem to develop these two characteristics while discouraging the third Class III variable, Achievement via independence (Ai). Thus, it is hypothesized that on the latter variable superior pupils in random grouped classrooms would achieve more.

Slow pupils in District A were also hypothesized to obtain higher Ac scores than their District R counterparts. No other between-treatment differences were hypothesized for the Class III variables.

It appeared likely that comparisons between pupils at different ability levels would reveal all three Class III measures to be related to ability classification in the Utah study.

Class IV measures are described by the test author as those dealing with intellectual and interest modes. A working hypothesis was developed for only one of the three Class IV measures, Flexibility. It appeared that the District R classroom environment with enrichment would tend to develop flexibility to a greater degree at all ability levels than would the ability grouped classroom with acceleration.

## PREVIOUS RESEARCH RELATED TO PERSONALITY AND GROUPING

### Introduction

The controversy over personality outcomes of ability grouping is to a large extent based on opinion and philosophical orientation. As a matter of fact the philosophy and attitude of teachers, administrators, and community do tremendously affect school programs of whatever nature.

One of the important questions in the conflict is, "Which kind of grouping is more democratic?" Distortions of the ideas that "All men are created equal," and "Equal treatment for all," are still interfering with objective, realistic evaluation of the problems involved in meeting, accepting, and handling individual differences in the educational system. A "set" in any particular orientation tends to interfere with

Table 12-1

Personality Variables Measured on the *CPI*[a]

| | | |
|---|---|---|
| Class I. Measures of Poise, Ascendancy, and Self-Assurance | | |
| 1. | Do | Dominance |
| 2. | Cs | Capacity for Status |
| 3. | Sy | Sociability |
| 4. | Sp | Social Presence |
| 5. | Sa | Self-acceptance |
| 6. | Wb | Sense of Well-being |
| Class II. Measures of Socialization, Maturity, and Responsibility | | |
| 7. | Re | Responsibility |
| 8. | So | Socialization |
| 9. | Sc | Self-control |
| 10. | To | Tolerance |
| 11. | Gi | Good impression |
| 12. | Cm | Communality |
| Class III. Measures of Achievement Potential and Intellectual Efficiency | | |
| 13. | Ac | Achievement via conformance |
| 14. | Ai | Achievement via independence |
| 15. | Ie | Intellectual efficiency |
| Class IV. Measures of Intellectual and Interest Modes | | |
| 16. | Py | Psychological-mindedness |
| 17. | Fx | Flexibility |
| 18. | Fe | Femininity |

---

[a] Taken from Gough, H. G., *Manual for the California Psychological Inventory*. Palo Alto, California: Consulting Psychologists Press, 1957, Page 7.

open-minded evaluations and interpretations of findings. Selective perception is potent. The "set" of the citizenry, as well as the "set" of educators and researchers, is involved. Since the citizenry establish the goals of education, send their children for that education, and pay the costs, probably they and their "sets" have not received the attention they deserve in evaluation of outcomes.

The "education of the layman" is underway, as indicated above. Probably the demand of parents of mentally and otherwise handicapped children for special education for their children has been one of the most influential factors dramatizing the fact of individual differences and the need for differential instruction. The recently accentuated emphasis on the problems, education, and society's need of the gifted has added to this awareness. However, in spite of the growing awareness of both the extent and complexity of the problem the controversy still rages.

## Ability Grouping

One of the most severe critics of ability grouping on philosophical grounds is Bruno Bettelheim, who said, "The conscious concern for the gifted child is an unconscious desire to create segregation based on intelligence" (Bettelheim, 1958, p.263). He warns that segregated education would ultimately lead to something similar to Naziism. He contends further that the gifted child develops far better in a heterogeneous setting. In reporting his observations of a group that had been in a "gifted program" and were later returned to a good public school he said:

> They began to reflect spontaneously on many problems . . . Not having to pay rapt attention to understand the point the teacher is making, the gifted student may have time to speculate on whether the teacher's point is worth making at all--a question he is far too busy to ponder in a higher accelerated program. (p. 267)

The "gifted class" was not described by Bettelheim, but judging from the quotation above, it must have been of a very different nature from those in Drews' study (1962) where the superior groups were practically self-directing.

Hamalainen (1950), in an opinion article, elaborated the social values of heterogeneous grouping, and concluded that for the sake of children's wholesome social development, heterogeneous grouping is the only method that should be used. In an opinion article Tonsor (1953) reached the same conclusion, adding that slow children resent being placed in slow groups. He claimed the heterogeneous group more nearly approaches the everyday life situation in which the individual will later find himself, and therefore furnishes a better setting for learning democratic processes.

Gowen (1955), in another opinion article, took the opposite point of view, contending that homogeneous grouping is democratic because in adult life, the family, work, and play, groups are relatively homogeneous. She claimed that a relatively homogeneous group is the only kind of group in which a student can practice and leran democratic processes. Her contention was that in a heterogeneous situation the child is more apt to fail to learn social empathy and social skills.

The articles cited above give the reader some insight into the emotions that are aroused among behavioral scientists when the topic of ability grouping and its effect upon personality and social development is discussed. These opinions, however, provide no real evidence concerning the relative effects of different grouping treatments upon the development of the pupil. Evidence regarding this question is very limited.

In an individual interview study, Luchins and Luchins (1948) questioned 190 fourth, fifth, and sixth grade children about their attitudes toward the ability grouping practiced in their school. All pupils were aware of the significance of the superscript, "1" (bright), "2" (dull), "3" (average) attached to his class. Practically all in the "1" class students preferred their class and would choose to remain in it even if they did not like the teacher. Among the reasons for choosing the "1" class were: they belonged there, they would learn more, they did not want to waste time, it was an honor, "2" class would be boring, they did not want to mingle with or be educated with "dumbbells." When asked which class they would rather be in, children in the average and dull classes chose "1" class because there would be stigma attached to being in "2" class which "has a reputation you can't get rid of," their parents would be proud of them, the teacher would not be so strict, they would have more fun, they would learn more, and besides it paid to take the chance since they could always go back to "2" class. The few students in the bright and average classes who selected the "2" class did so because they might get better grades or they could be more outstanding. Those in the "2" class who selected to remain said it was best for them, they would be out of place in "1" class, and they wouldn't be allowed in "1" class anyway, so why should they pick it? Luchins and Luchins concluded that the prestige value of "1" class membership was its chief attraction, that it caused the able members to become snobbish, that the grouping situation in the school caused stigma to be attached to the "2" class membership, and that homogeneous grouping created a caste system in the school.

Jackson (1943), principal of an elementary school, found that when slow children were grouped together with opportunities for success and good grades, truancy almost disappeared and behavior problems were inconsequential. His was an observational report, not a statistical one.

Marsh (1955), principal of a junior high school, in another observer report on ability grouping found that classroom discipline was reduced to the lowest level in the history of the school. The program involved teacher participation in planning and in multiple screening, using extensive teacher ratings (grades, standardized tests, and opinion); pupil choice of electives, friends and teachers; and careful class scheduling. Teachers felt they were doing a better job and pupil needs were provided for much better, resulting in higher morale. Most of the children were challenged to put forth more effort.

Kvaraceus (1938) reported on a school that found its achievement level to be the lowest in its district. In an effort to remedy the situation a program of partial segregation was initiated. The homeroom and social studies classes were heterogeneously grouped, and other subject area classes were ability grouped in three levels according to achievement tests and teacher judgment. All teachers worked with all three achievement levels during the day, thus enabling them to keep their perspective. During the experimental year the academic level of the school was raised from lowest to fifth from the lowest in the district, and disciplinary problems were reduced.

Riley (1956), reporting on a four level grouping study, said that the children in the slower classes were able to achieve, and were therefore happy, and that the social development of the slower children in their ability to hold class offices was significant.

Mann (1957), found in a program of partial segregation (Colfax plan) that children in the fourth, fifth, and sixth grades, both gifted and average, chose and rejected other children of their own group significantly more often than they chose and rejected children from the other ability levels. His conclusion was that placing gifted and average children together does not necessarily stimulate friendships between the two groups. Putting it a different way, even though there was partial heterogeneous grouping, there still was a tendency for the "birds of a feather" process to function. He made no mention of the fact that both the average and gifted spent the whole day together, and only a part of the day in the heterogeneous situation.

Mathias (1959) found, in a *Rorschach* and *TAT* study where he measured aggression and depression, that the slow and bright children had significantly more aggression responses (beyond the .01 level of significance) than the average in heterogeneous classes. He concluded this was due to frustration they experienced in a school program aimed primarily at the average pupil.

Simpson and Martinson (1961) administered the *California Psychological Inventory* to their special class groups and control groups at the eighth grade and high school levels. Subjects involved in the eighth grade phase of the Simpson and Martinson study included 58 boys and 57 girls in the special classes and 36 boys and 20 girls in the control group. It will be recalled that this study was concerned with superior pupils only. The special class program was similar to ability grouping except that it involved superior pupils only. The *CPI* was administered to the special class and control groups at the beginning and end of the one-year study and gains in the subtest scores were tested for significance. Eighth grade boys in the special classes made significant gains in *Self-Acceptance* and *Social Presence*. Eighth grade boys in the control group made significant gains on *Dominance*, *Well-Being*, *Socialization*, *Communality*, *Social Presence*, and *Tolerance*. Eighth grade girls in special classes made significant gains in *Tolerance*, *Capacity for Status*, *Social Presence*, *Self-Acceptance*, and *Flexibility*; while the girls' control group made no significant gains between the initial and final testing.

The high school phase of the study included 109 boys and 73 girls in the special class groups with 48 boys and 34 girls in the control groups. Gifted high school boys in the experimental programs made significant gains in *Social Presence*, *Self-Acceptance*, *Tolerance*, *Achievement via Independence*, *Flexibility*, and *Capacity for Status*. Boys in the high school control groups made significant gains in *Dominance*, *Self-Acceptance*, and *Capacity for Status*. Gifted high school girls in the special classes made significant gains in *Self-Acceptance*, *Tolerance*, *Achievement via Independence*, *Flexibility*, *Dominance and Social Presence*. Significant gains were made by the high school girls' control groups in *Social Presence*, *Self-Acceptance*, *Tolerance*, and *Achievement via Independence*.

The special class groups made significant gains in 19 instances and signigicant losses in three instances on the *CPI*. The control group pupils made significant gains in nine instances and significant losses in eight instances on the *CPI*.

This study did not present direct statistical comparisons to determine whether the experimental groups made significantly greater gains than the control groups. Examination of the eighth grade data available for boys would indicate, however, that the special class group probably made significantly greater gains on *Self-Acceptance* than the control group. Similar comparisons for the eighth grade girls appear to indicate that the special class girls made a significantly greater relative gain in *Self-Acceptance* and *Flexibility* than the control group.

The data available for the high school boys in the special class and control groups would indicate that the special class sample made a significantly greater gain in *Social Presence* and *Tolerance*.

High school girls in the control group appear to have made significantly greater gains than the special class group in Social Presence. Girls in the special class and control groups did not appear to be significantly different on any of the other CPI variables.

## PROCEDURE

### Personality Measures

#### California Test of Personality (CTP)

Parts of the CTP elementary level, were administered to Sample IV pupils during their fifth grade year. This test is a fairly direct self-report type of personality inventory. Two major areas are measured: Personal Adjustment and Social Adjustment. The Personal Adjustment area includes six sub-areas of which Sense the Personal Worth, Feeling of Belonging, and Withdrawing Tendencies were measured in this research. The Social Adjustment area also contains six sub-tests of which Anti-social Tendencies and School Relations were measured in this study. Sense of Personal Worth appears to be essentially a measure of self-concept and results on this section are treated in Chapter 10. School Relations appears to be most closely related to school attitudes and is treated in Chapter 8. Reliability coefficients for the subtests of the California Test of Personality, elementary level, that were employed in this research may be found in Table 12-2. Although these subtests are short, it may be seen that all subtests scores have sufficient reliability for research purposes.

As is the case with most direct self-report measures, the items themselves present some evidence of content validity. Such measures, however, are dependent upon the accuracy of the subject's insight and his willingness to provide honest answers. In the Utah study this measure was administered in a permissive atmosphere and pupils were given no motives for falsifying of faking their responses. In the more important area of concurrent validity, little evidence is available on this measure. Only one study directly pertinent to concurrent validity of the elementary form of the CTP is reported in the test manual. Taylor and Combs (1952) demonstrated that sixth grade pupils scoring in the upper 50 per cent on total adjustment were more willing to check self-damaging statements than the more poorly adjusted 50 per cent. This difference was statistically significant and indicates that the test has some concurrent validity as a measure of total adjustment. No concurrent validity data, however, have been reported for the various subtests.

#### Inventory of Factors GAMIN

The Guilford-Martin Inventory of Factors GAMIN provides scores on five personality variables developed through factor analyses by Guilford and his associates (Guilford and Guilford, 1939; Mosier, 1937). Two of the variables on this measure purport to test characteristics that educators have often suggested were influenced by ability grouping. Factor A, which measures ascendancy in social situations as opposed to submissiveness, seemed likely to be influenced by the differences in the social structure of ability grouped and random grouped classrooms. Factor I, which measures lack of inferiority feeling or self-confidence, offered an opportunity to test the often heard contention that ability grouping causes inferiority feelings and lack of self-confidence among slow pupils. Items measuring these two factors were administered to Sample VI pupils during their seventh grade year, Sample VII pupils

Table 12-2

Reliability Coefficients of the California Test of Personality[a]

| Components | Reliability |
|---|---|
| Sense of Personal Worth | .79 |
| Feeling of Belonging | .77 |
| Withdrawing Tendencies | .83 |
| Anti-social Tendencies | .77 |
| School Relations | .78 |

[a] From Thorpe, L. P. Clark, W. W., and Tiegs, E. W., Manual, California Test of Personality, 1953 Revision. Los Angeles, California, Test Bureau, 1953, p. 4.

during their eighth grade year, and Sample VIII pupils during their ninth grade year. In each case, District A pupils had been exposed to the ability grouping treatment for a little more than one and one-half school years. Although this measure is usually not administered below grade nine, the reading level and content of the items appear generally appropriate for pupils as low as grade seven. Factor A contains 38 items and the split-half reliability of this measure as reported in the test manual is .88. Factor I contains 49 items and has a split-half reliability of .89 (Guilford and Martin, 1948).

### Objective-Analytic Personality Test Factors

At the time that the test battery for this research project was being assembled, the preliminary forms of the Objective-Analytic Personality Test Batteries by R. B. Cattell and his associates had been available to psychologists for about three years. These measures, although in an experimental form, appeared to offer certain advantages over both the group administered personality inventories and the projective techniques. As Cattell (1955, p. 1) points out in the test handbook:

> Previously available tests--such as questionnaires and open-ended projective tests--have been non-objective, in the sense that (a) the subject can see how to "doctor" his responses to give whatever impression he desires, and (b) the psychologist, in "intuitive" scoring adds his own personal equation to that of the subject, and (c) there is no clear and proven theory as to the identity or the trait of psychological state being measured.
>
> In these three respects the Objective-Analytic batteries present major advances; they measure behavior in actual situations, instead of introspective self-assessments; they are scored by keys through which all psychometrists get the same answers; and they measure functional unities in personality structure established by repeated factor analytic and experimental investigations.

The factors forming the basis of this battery emerged from a series of factor analysis studies by Cattell and a number of other researchers. This research has resulted in the identification of 18 personality factors, the first 12 of which Cattell considers sufficiently invariant and clearly proved by research to be used with confidence (1955). Fewer factors have been identified for children, and the factors that have emerged for children are based on fewer studies than those for adults. Cattell has provided a number of individual and group tests that can be used to measure each personality factor in his battery. Children's tests often differ from the adult tests in level of sophistication and in some cases in factor loadings.

Two factors, taken from the 12 Cattell considers best established, were selected for use in this research. The factors were UI-16 Competent Assertiveness, and UI-24, Anxiety to Achieve. Cattell advises the use of at least six tests to obtain each factor score. As the research situation in the Utah study did not permit the use of individual tests, six group tests were selected for each of the two factors. Tests selected for Factor UI-16 Competent Assertiveness were: G-6, G-20, G-27, G-42, G-44a, and G-45. For Factor UI-24, Anxiety to Achieve, tests G-26, G-29, G-31, G-38, G-41, and G-47 were selected. These tests employ a wide variety of approaches, but generally have no obvious relationship to the factor being measured and therefore are not fakable. Test G-20, for example, is made up of a set of eight unstructured drawings and calls for responses such as are often asked for on projective tests. Test G-45 is made up of pairs of straight lines and requires the subject to judge which line is longer or whether the two are equal. Test G-47 contains a number of short paragraphs involving three characters and asks the pupil to indicate his preference or liking for the characters. The interested reader is referred to the Handbook for the Objective-Analytic Personality Test Batteries (Cattell, 1955) and to research published on this battery in recent years.

A preliminary trial of the above measures indicated that the test instructions were generally inadequate for children. Therefore, most of the instructions were rewritten and more examples were given in order to make them easier to understand. The basic meanings of the instructions, however, were not changed.

## The California Psychological Inventory

The California Psychological Inventory (CPI) contains 480 items and yields scores on 18 personality variables, which will be found listed in Table 12-1. It will be recalled that the 18 scales are classified by the test author into four broad areas or classes. Twelve of the CPI scales were developed using the "empirical technique." In this technique, a criterion dimension is first defined and items are then developed which appear to be psychologically relevant to this dimension. These items are then administered to a criterion group and item-analyzed to identify those which discriminate between individuals who are high and low on the criterion dimension. This method builds contruct validity into the scale to the degree that the test developer employs psychological constructs in his initial preparation of the items. Concurrent validity is established through the initial item-analysis and through cross validational studies. Cross validational studies for each of the 18 scales of the CPI are summarized in the test manual (Gough, 1957). The principal limitations of this method of scale development are first, those imposed by the psychological sophistication of the test author, and secondly, limitations imposed by the availability of satisfactory criterion groups. The CPI items generally reflect a high level of psychological sophistication thus overcoming the first limitation and are somewhat less direct than items found on many personality inventories. The degree to which the second limitation is met varies with the different scales; some having much more adequate criterion groups than others.

In the case of four scales (Sp, Sa, Sc, and Fx) satisfactory criterion samples could not be identified and the scales were developed through internal consistency analysis. Since the development of these four scales, however, concurrent validity studies have been carried out on all of them and the validities obtained are comparable to those for the other scales.

Reliability coefficients on the 18 scales generally range around .70. Test-retest reliability coefficients as given in the test manual are shown in Table 12-3. The Cm (Communality) and Py (Psychological-mindedness) scales have the lowest reliability for both the male and female high school samples. Lower reliability on the Cm scale is probably attributable to the fact that many subjects obtain very high scores on this scale, thus markedly skewing the distribution. The Py scale contains only 22 items and its lower reliability is probably due to its relative shortness. All in all, however, the reliability coefficients for the various scales of the CPI are more than adequate for research applications in which comparisons are to be made between groups rather than individuals.

Concurrent validities on the 18 scales are reported in a range from .26 to .83, with the majority yielding correlations in the .30 to .50 range with the concurrent criterion measures (Gough, 1957). The validity evidence for the CPI is substantial, although some of the validity coefficients are not as high as one might desire. When it is considered, however, that these coefficients are attenuated to a considerable degree by the reliability of both the scales and criterion measures and in many cases represent correlations with criterion measures that would not be expected to be highly reliable, the reported validities appear to be quite adequate for research purposes. Most of the concurrent validities are sufficiently high so that any marked influence of the grouping treatments upon these personality variables should be revealed in the analysis.

### Sample

Several personality tests and related measures were employed in the Utah study. The samples to whom the various measures were administered are shown in Table 12-4. It will be noted that most of these measures were administered during the junior high school years. The lack of suitable measures appropriate for administration at the elementary school level reduced the amount of data that could be collected at this level. On the other hand, little personality data were collected at the high school level because of the limited differences between the grouping treatments as applied at the high school level in Districts A and R. As a rule, all subjects in a given sample were administered the personality measures scheduled for use with that sample, excepting the cases who were not available for either the regular or make-up testing. The projective measures, however, because of the expense of administration and scoring, were administered only to subgroups within Sample IV.

## RESULTS

### California Test of Personality

Between-treatment comparisons on the Feeling of Belonging subscores of this measure (Table 12-5) showed District R subgroups to be consistently higher than comparable District A subgroups. These differences reached statistical significance in three of the six comparisons, one at each ability level. Comparisons between pupils of different ability levels revealed a consistent tendency for pupils of higher ability to obtain more favorable scores on this subscale in both districts. All 12 of the within-treatment comparisons (Table 12-6) revealed differences in this direction with seven of these differences reaching statistical significance. Between-sex comparisons for pupils at the same ability level in each district revealed only one significant difference, favoring slow girls in District A (Table 12-7).

These data suggest that pupils generally experienced a greater feeling of belonging in District R at all ability levels than in District A and also seem to indicate that superior pupils in both districts experience a greater feeling of belonging than pupils of lower ability level. In interpreting these findings, the reader should bear in mind that, as the California Test of Personality was administered to Sample IV only, the findings may be accepted with somewhat less confidence than would have been the case if consistent results had been obtained for several samples. Also, the absence

Table 12-3

Test-Retest Reliability Coefficients on the CPI[a]

| Scale | High School Females (N-125) | High School Males (N-101) |
|---|---|---|
| Do | .72 | .64 |
| Cs | .68 | .62 |
| Sy | .71 | .68 |
| Sp | .63 | .60 |
| Sa | .71 | .67 |
| Wb | .72 | .71 |
| Re | .73 | .65 |
| So | .69 | .65 |
| Sc | .68 | .75 |
| To | .61 | .71 |
| Gi | .68 | .69 |
| Cm | .44 | .38 |
| Ac | .73 | .60 |
| Ai | .57 | .63 |
| Ie | .77 | .74 |
| Py | .49 | .48 |
| Fx | .67 | .60 |
| Fe | .65 | .59 |

[a] Taken from Gough, H. G., Manual for the California Psychological Inventory. Palo Alto, California, Consulting Psychologists Press, 1957, p. 22.

Table 12-4

Samples Administered Personality and Related Measures in the Utah Study

| | Sample and Grade Level to Which Each Measure was Administered | | | | |
|---|---|---|---|---|---|
| Measure | Sample IV | Sample VI | Sample VII | Sample VIII | Sample IX |
| 1. California Test of Personality (five subtests used) | 5th | | | | |
| 2. Inventory of Factors GAMIN (two factors used, ascendancy and inferiority) | | 7th | 8th | 9th | |
| 3. Objective-Analytic Personality Test Factors | | | | | |
| UI-16 -- Competent Assertiveness | 5th | | | | |
| UI-24 -- Anxiety to Achieve | | 7th | | | |
| 4. California Psychological Inventory (all 18 scores obtained) | | 8th | 9th | | |
| 5. Individually administered TAT type measure | 5th (slow) | | | | |
| 6. Group administered TAT type measure | 6th | | | | |
| 7. Index of Adjustment and Values (self-concept) | 5th,6th<br>7th | 7th,8th<br>9th | 8th,9th | 9th<br>10th | 9th<br>11th |
| 8. SRA Junior and Youth Inventories (Problem check lists) | 5th | 7th | 8th | 9th | |

of concurrent validity evidence and the few items employed to measure each of the subscale areas would suggest that these results be interpreted with considerable caution.

## Withdrawing Tendencies

Comparisons between District A and District R pupils on the Withdrawing Tendencies subscale revealed only one statistically significant difference favoring average boys in District R. The non-significant differences on this subscale are all in the direction of higher scores for District R pupils thus presenting a rather consistent picture. Comparisons between pupils at different ability levels on this subscale revealed a pattern similar to that found on the Feeling of Belonging subscore; a rather consistent tendency for pupils of higher ability to obtain more favorable scores. All 12 of the within-treatment comparisons on this variable favored the group of higher ability with eight of these differences reaching statistical significance at the .05 level. There were no significant sex differences on this subscale of the California Test of Personality.

## Anti-Social Tendencies

Between-treatment comparisons on this subscale revealed four significant differences in the six comparisons made, all favoring the District R groups. The same consistent tendency for superior pupils to obtain higher scores on the within-treatment comparisons that was found on the previously discussed subscales of this inventory is also present in the Anti-Social Tendencies scores. Of the 12 within-treatment comparisons, seven were statistically significant, all favoring the group of higher ability. Two significant sex differences were also found on this score. In both districts, superior girls obtained significantly more favorable scores than superior boys. Differences between average and slow boys and girls consistently reflected a tendency for girls to obtain more favorable scores.

## Total Adjustment

Total adjustment was estimated on the California Test of Personality by combining scores on the five subtests administered in the Utah study. Between-treatment comparisons on total adjustment showed that District R pupils obtained more favorable scores in all six comparisons with five of these reaching statistical significance. Because of the limitations in both the sample and the measure employed, these findings are tentative but provide some evidence to suggest that pupils at all ability levels tend to report more favorable adjustments in random grouped classrooms. Within-treatment comparisons of the total adjustment scores showed a consistent pattern in both districts for pupils of higher ability to obtain more favorable scores.[1]

---

[1] The reader is reminded that the total adjustment score used in the Utah study is not comparable to total adjustment as reported in the CTP manual, the latter being based on all 12 of the CTP subtests.

## Table 12-5

Sample IV: Between-Treatment Differences on California Test of Personality Variables

| Group | $N_A$ | $N_R$ | Sense of Personal Worth $M_A$ | $M_R$ | t | Feelings of Belonging $M_A$ | $M_R$ | t | Withdrawing Tendencies $M_A$ | $M_R$ | t | Personal Adjustment Subtotal $M_A$ | $M_R$ | t |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Superior Boys, District A vs. R | 58 | 116 | 8.59 | 9.31 | 1.99 | 9.77 | 10.18 | N.S.[b] | 7.91 | 8.34 | N.S. | 26.29 | 27.83 | N.S. |
| Superior Girls, District A vs. R | 74 | 131 | 9.51 | 10.10 | 1.97 | 9.80 | 10.32 | 2.16 | 7.77 | 8.44 | N.S. | 27.23 | 28.89 | 1.99 |
| Average Boys, District A vs. R | 124 | 116 | 8.07 | 9.12 | 3.51 | 9.25 | 9.94 | 2.72 | 7.00 | 8.03 | 2.69 | 24.34 | 26.99 | 3.41 |
| Average Girls, District A vs. R | 96 | 112 | 8.88 | 9.39 | A.S.[a] | 9.31 | 9.62 | N.S. | 6.77 | 7.46 | N.S. | 24.98 | 26.37 | N.S. |
| Slow Boys, District A vs. R | 57 | 76 | 7.32 | 8.07 | A.S. | 7.49 | 8.91 | 3.19 | 6.30 | 6.99 | N.S. | 21.14 | 23.96 | 2.45 |
| Slow Girls, District A vs. R | 32 | 38 | 8.28 | 9.05 | N.S. | 8.78 | 9.16 | N.S. | 6.19 | 6.97 | N.S. | 23.31 | 25.08 | N.S. |

| Group | Anti-Social Tendencies | | | School Relations | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys, District A vs. R | 8.21 | 9.08 | 2.07 | 8.34 | 9.14 | 2.11 | 42.71 | 46.03 | 2.09 |
| Superior Girls, District A vs. R | 9.35 | 9.75 | N.S. | 9.35 | 10.01 | 2.11 | 45.81 | 48.65 | 2.18 |
| Average Boys, District A vs. R | 7.55 | 8.51 | 2.65 | 8.06 | 8.68 | N.S. | 39.95 | 44.19 | 3.20 |
| Average Girls, District A vs. R | 8.21 | 9.00 | 2.17 | 8.46 | 9.06 | N.S. | 41.66 | 44.42 | 1.96 |
| Slow Boys, District A vs. R | 6.40 | 8.05 | 3.32 | 7.09 | 8.07 | 2.04 | 34.65 | 40.09 | 2.89 |
| Slow Girls, District A vs. R | 7.59 | 8.68 | N.S. | 7.91 | 8.74 | N.S. | 38.81 | 42.50 | N.S. |

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

Table 12-6

Sample IV--Within-Treatment Differences on California Test of Personality Variables

| Group | $N_1$ | $N_2$ | Sense of Personal Worth $M_1$ | $M_2$ | t | Feeling of Belonging $M_1$ | $M_2$ | t | Withdrawing Tendency $M_1$ | $M_2$ | t | Personal Adjustment Subtotal $M_1$ | $M_2$ | t |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| District A | | | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 58 | 124 | 8.59 | 8.07 | N.S.[b] | 9.76 | 9.25 | A.S.[a] | 7.91 | 7.00 | 1.98 | 26.29 | 24.34 | 2.08 |
| (1) Superior Boys vs.<br>(2) Slow Boys | 58 | 57 | 8.59 | 7.32 | 2.91 | 9.76 | 7.49 | 5.09 | 7.91 | 6.30 | 2.81 | 26.29 | 21.14 | 4.34 |
| (1) Average Boys vs.<br>(2) Slow Boys | 124 | 57 | 8.07 | 7.32 | 1.98 | 9.25 | 7.49 | 5.02 | 7.00 | 6.30 | N.S. | 24.34 | 21.14 | 3.28 |
| (1) Superior Girls vs.<br>(2) Average Girls | 74 | 96 | 9.51 | 8.89 | A.S. | 9.80 | 9.31 | N.S. | 7.76 | 6.77 | 2.06 | 27.23 | 24.98 | 2.39 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 74 | 32 | 9.51 | 8.28 | 2.92 | 9.80 | 8.78 | 2.76 | 7.76 | 6.19 | N.S. | 27.23 | 23.31 | N.S. |
| (1) Average Girls vs.<br>(2) Slow Girls | 96 | 32 | 8.89 | 8.28 | N.S. | 9.31 | 8.78 | N.S. | 6.77 | 6.19 | N.S. | 24.98 | 23.31 | N.S. |
| District R | | | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 116 | 116 | 9.31 | 9.11 | N.S. | 10.18 | 9.94 | N.S. | 8.34 | 8.03 | N.S. | 27.83 | 27.00 | N.S. |
| (1) Superior Boys vs.<br>(2) Slow Boys | 116 | 76 | 9.31 | 8.06 | 3.69 | 10.18 | 8.91 | 4.29 | 8.34 | 6.99 | 3.01 | 27.83 | 23.96 | 4.25 |
| (1) Average Boys vs.<br>(2) Slow Boys | 116 | 76 | 9.11 | 8.06 | 3.15 | 9.94 | 8.91 | 3.21 | 8.03 | 6.99 | 2.24 | 27.00 | 23.96 | 3.23 |
| (1) Superior Girls vs.<br>(2) Average Girls | 131 | 112 | 10.10 | 9.39 | 2.58 | 10.33 | 9.62 | 3.13 | 8.44 | 7.45 | 2.60 | 28.88 | 26.37 | 3.32 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 131 | 38 | 10.10 | 9.05 | 2.70 | 10.33 | 9.16 | 3.54 | 8.44 | 6.97 | 2.83 | 28.88 | 25.08 | 3.58 |
| (1) Average Girls vs.<br>(2) Slow Girls | 112 | 38 | 9.39 | 9.05 | N.S. | 9.62 | 9.16 | N.S. | 7.45 | 6.97 | N.S. | 26.37 | 25.08 | N.S. |

| Group | $N_1$ | $N_2$ | Anti-Social Tendencies $M_1$ | Anti-Social Tendencies $M_2$ | Anti-Social Tendencies t | School Relations $M_1$ | School Relations $M_2$ | School Relations t | Total $M_1$ | Total $M_2$ | Total t |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **District A** | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 58 | 124 | 8.21 | 7.55 | N.S. | 8.34 | 8.06 | N.S. | 42.71 | 39.95 | A.S. |
| (1) Superior Boys vs.<br>(2) Slow Boys | 58 | 57 | 8.21 | 6.40 | 3.45 | 8.34 | 7.09 | 2.42 | 42.71 | 34.65 | 4.09 |
| (1) Average Boys vs.<br>(2) Slow Boys | 124 | 57 | 7.55 | 6.40 | 2.67 | 8.06 | 7.09 | 2.33 | 39.95 | 34.65 | 3.34 |
| (1) Superior Girls vs.<br>(2) Average Girls | 74 | 96 | 9.35 | 8.21 | 2.89 | 9.35 | 8.46 | 2.35 | 45.81 | 41.66 | 2.71 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 74 | 32 | 9.35 | 7.59 | 3.25 | 9.35 | 7.91 | 2.77 | 45.81 | 38.81 | 3.46 |
| (1) Average Girls vs.<br>(2) Slow Girls | 96 | 32 | 8.21 | 7.59 | N.S. | 8.46 | 7.91 | N.S. | 41.66 | 38.81 | N.S. |
| **District R** | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 116 | 116 | 9.08 | 8.51 | N.S. | 9.14 | 8.68 | N.S. | 46.03 | 44.19 | N.S. |
| (1) Superior Boys vs.<br>(2) Slow Boys | 116 | 76 | 9.08 | 8.05 | 2.61 | 9.14 | 8.07 | 3.10 | 46.03 | 40.09 | 3.98 |
| (1) Average Boys vs.<br>(2) Slow Boys | 116 | 76 | 8.51 | 8.05 | N.S. | 8.68 | 8.07 | A.S. | 44.19 | 40.09 | 2.55 |
| (1) Superior Girls vs.<br>(2) Average Girls | 131 | 112 | 9.75 | 9.00 | 2.60 | 10.01 | 9.06 | 3.31 | 48.65 | 44.42 | 3.53 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 131 | 38 | 9.75 | 8.68 | 2.89 | 10.01 | 8.74 | 3.39 | 48.65 | 42.50 | 3.79 |
| (1) Average Girls vs.<br>(2) Slow Girls | 112 | 38 | 9.00 | 8.68 | N.S. | 9.06 | 8.74 | N.S. | 44.42 | 42.50 | N.S. |

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

Table 12-7

Sample IV - Sex Differences on California Test of Personality Variables

| | | | Sense of Personal Worth | | | Feelings of Belonging | | | Withdrawing Tendencies | | | Personal Adjustment Subtotal | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | $N_B$ | $N_G$ | $M_B$ | $M_G$ | t | $M_B$ | $M_G$ | t | $M_B$ | $M_G$ | t | $M_B$ | $M_G$ | t |
| **District A** | | | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | 58 | 74 | 8.59 | 9.51 | 2.42 | 9.78 | 9.79 | N.S.[b] | 7.91 | 7.76 | N.S. | 26.29 | 27.23 | N.S. |
| Average Boys vs. Average Girls | 124 | 96 | 8.07 | 8.88 | 2.60 | 9.25 | 9.31 | N.S. | 7.00 | 6.77 | N.S. | 24.34 | 24.98 | N.S. |
| Slow Boys vs. Slow Girls | 57 | 32 | 7.32 | 8.28 | 2.01 | 7.49 | 8.78 | 2.37 | 6.30 | 6.19 | N.S. | 21.14 | 23.31 | N.S. |
| **District R** | | | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | 116 | 131 | 9.31 | 10.09 | 2.90 | 10.18 | 10.32 | N.S. | 8.34 | 8.44 | N.S. | 27.83 | 28.88 | N.S. |
| Average Boys vs. Average Girls | 116 | 112 | 9.11 | 9.39 | N.S. | 9.94 | 9.62 | N.S. | 8.03 | 7.46 | N.S. | 26.99 | 26.37 | N.S. |
| Slow Boys vs. Slow Girls | 76 | 38 | 8.06 | 9.05 | 2.12 | 8.90 | 9.16 | N.S. | 6.99 | 6.97 | N.S. | 23.96 | 25.08 | N.S. |

| Group | Anti-Social Tendencies | | | School Relations | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | $M_B$ | $M_G$ | t | $M_B$ | $M_G$ | t | $M_B$ | $M_G$ | t |
| District A | | | | | | | | | |
| Superior Boys vs. Superior Girls | 8.21 | 9.35 | 2.65 | 8.35 | 9.35 | 2.34 | 42.71 | 45.81 | A.S.[a] |
| Average Boys vs. Average Girls | 7.55 | 8.21 | A.S. | 8.06 | 8.46 | N.S. | 39.95 | 41.66 | N.S. |
| Slow Boys vs. Slow Girls | 6.40 | 7.59 | A.S. | 7.09 | 7.91 | N.S. | 34.65 | 38.81 | A.S. |
| District R | | | | | | | | | |
| Superior Boys vs. Superior Girls | 9.08 | 9.75 | 2.36 | 9.14 | 10.01 | 3.23 | 46.03 | 48.65 | 2.24 |
| Average Boys vs. Average Girls | 8.51 | 9.00 | N.S. | 8.68 | 9.06 | N.S. | 44.19 | 44.42 | A.S. |
| Slow Boys vs. Slow Girls | 8.05 | 8.68 | N.S. | 8.06 | 8.74 | N.S. | 40.09 | 42.50 | N.S. |

---

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

## Inventory of Factors GAMIN

Two of the personality factors from this inventory, ascendancy and inferiority, were hypothesized to be influenced by differences in the grouping treatments. Scores on these factors were obtained for Samples VI, VII, and VIII during the second year of the study.

### Ascendancy

Between-treatment differences on the Ascendancy score are summarized in Table 12-8. Of the 18 comparisons made between pupils in District A and District R on this factor, only two reached statistical significance. Both of the significant differences were found in Sample IV, one favoring superior District R boys and the other favoring slow District R girls. No consistent trends were apparent among the mean scores of comparable District A and R groups. Most of the differences between comparable groups were very small with 11 favoring District R groups and seven favoring District A groups. The lack of any consistent pattern in the differences obtained on this score plus the fact that only two significant differences were obtained from 18 comparisons seems to suggest that no true generalizable differences in the Ascendancy score occurred as a result of differences between the two grouping situations.

Within-treatment differences on the Ascendancy score (Table 12-9) showed a slight but generally non-significant trend in both districts for pupils of higher ability to obtain higher scores. In District A, only two significant differences occurred in the 18 comparisons between pupils of different ability level. These two significant differences were both found in Sample VI and reflect the very low Ascendancy score obtained by slow girls in this sample. It is possible that this low score reflected the experience of these girls with ability grouping at the sixth grade level, where the grouping situation is somewhat different than found at the junior high school level. On the other hand, the difference may merely reflect some unique characteristic of this sample. It will be noted that the slow girls in Sample VI only included 23 cases for this analysis. Of the 18 within-treatment comparisons made for District R pupils, two reached statistical significance. District R superior boys in Sample VIII obtained higher scores than average or slow boys in this sample. It will be noted that superior boys in Sample VI also attained somewhat higher Ascendancy scores than average and slow boys, with both of these differences approaching significance. This same trend is reflected in Sample VII although differences for this sample are somewhat smaller. Thus, of the six comparisons between boys of superior and average or slow ability in District R two reached statistical significance, and two approached significance. These data would suggest that there is a fairly consistent tendency for superior boys in random grouped classrooms to obtain higher Ascendancy scores than average or slow boys. No significant differences were found between average and slow boys in District R or in any of the comparisons between District R girls at different ability levels, suggesting that for these groups ability level is not significantly related to ascendancy.

Sex difference comparisons on the Ascendancy score reflect a consistent tendency for boys to obtain more favorable scores on this factor than girls (see Table 12-11). In Sample VI, five of the six comparisons between the sexes were statistically significant. One difference was significant for Sample VII, and three for Sample VIII, all favoring the boy's groups. Differences between the sexes were most pronounced for superior pupils with five of the six comparisons between superior boys

Table 12-8

Between-Treatment Differences on the GAMIN Ascendancy and Inferiority Scores

| Group | Mean A | Mean R | t | Mean A | Mean R | t | Mean A | Mean R | t |
|---|---|---|---|---|---|---|---|---|---|
| | | | | **Ascendancy** | | | | | |
| | Sample VI | | | Sample VII | | | Sample VIII | | |
| Superior Boys--Dist. A vs. R | 16.05 | 18.05 | 1.96 | 16.35 | 16.07 | NS[a] | 17.23 | 17.54 | NS |
| Superior Girls--Dist. A vs. R | 13.81 | 13.94 | NS | 14.32 | 15.46 | NS | 13.92 | 14.89 | NS |
| Average Boys--Dist. A vs. R | 15.82 | 16.35 | NS | 15.71 | 15.05 | NS | 15.99 | 14.54 | NS |
| Average Girls--Dist. A vs. R | 13.65 | 14.47 | NS | 14.28 | 13.81 | NS | 14.18 | 13.51 | NS |
| Slow Boys--Dist. A vs. R | 16.03 | 16.05 | NS | 14.83 | 14.59 | NS | 15.43 | 14.10 | NS |
| Slow Girls--Dist. A vs. R | 10.09 | 13.89 | 3.24 | 13.17 | 15.03 | NS | 13.11 | 14.57 | NS |
| | | | | **Inferiority** | | | | | |
| | Sample VI | | | Sample VII | | | Sample VIII | | |
| Superior Boys--Dist. A vs. R | 24.51 | 25.29 | NS | 26.16 | 24.57 | NS | 25.85 | 25.02 | NS |
| Superior Girls--Dist. A vs. R | 24.18 | 24.80 | NS | 22.80 | 26.46 | 2.21 | 23.31 | 24.92 | NS |
| Average Boys--Dist. A vs. R | 22.59 | 22.96 | NS | 23.14 | 22.46 | NS | 22.84 | 22.71 | NS |
| Average Girls--Dist. A vs. R | 22.09 | 22.76 | NS | 21.06 | 20.85 | NS | 22.37 | 19.93 | NS |
| Slow Boys--Dist. A vs. R | 21.29 | 22.62 | NS | 21.28 | 21.05 | NS | 26.57 | 21.07 | 2.50 |
| Slow Girls--Dist. A vs. R | 13.70 | 20.69 | NS | 17.50 | 22.70 | 2.08 | 22.65 | 20.11 | NS |

[a] N.S. indicates difference not significant.

Table 12-9

Within-Treatment Differences on the GAMIN Ascendancy Score

| Group | Sample VI | | | | | Sample VII | | | | | Sample VIII | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $N_1$ | $N_2$ | $M_1$ | $M_2$ | t | $N_1$ | $N_2$ | $M_1$ | $M_2$ | t | $N_1$ | $N_2$ | $M_1$ | $M_2$ | t |
| **District A** | | | | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 65 | 172 | 16.05 | 15.82 | NS[b] | 49 | 138 | 16.35 | 15.71 | NS | 47 | 123 | 17.23 | 15.99 | NS |
| (1) Superior Boys vs.<br>(2) Slow Boys | 65 | 35 | 16.05 | 16.03 | NS | 49 | 36 | 16.35 | 14.83 | NS | 47 | 28 | 17.23 | 15.43 | NS |
| (1) Average Boys vs.<br>(2) Slow Boys | 172 | 35 | 15.82 | 16.03 | NS | 138 | 36 | 15.71 | 14.83 | NS | 123 | 28 | 15.99 | 15.43 | NS |
| (1) Superior Girls vs.<br>(2) Average Girls | 81 | 175 | 13.81 | 13.65 | NS | 60 | 116 | 14.32 | 14.28 | NS | 85 | 31 | 13.92 | 14.18 | NS |
| (1) Superior Girls vs.<br>(2) Slow Girls | 81 | 23 | 13.81 | 10.09 | 3.26 | 60 | 18 | 14.32 | 13.17 | NS | 85 | 40 | 13.92 | 13.12 | NS |
| (1) Average Girls vs.<br>(2) Slow Girls | 175 | 23 | 13.65 | 10.09 | 2.29 | 116 | 18 | 14.28 | 13.17 | NS | 31 | 40 | 14.18 | 13.12 | NS |
| **District R** | | | | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 58 | 92 | 18.05 | 16.35 | AS[a] | 79 | 57 | 16.07 | 15.05 | NS | 62 | 100 | 17.54 | 14.55 | 2.27 |
| (1) Superior Boys vs.<br>(2) Slow Boys | 58 | 58 | 18.05 | 16.05 | AS | 79 | 39 | 16.07 | 14.59 | NS | 62 | 26 | 17.54 | 14.10 | 2.93 |
| (1) Average Boys vs.<br>(2) Slow Boys | 92 | 58 | 16.35 | 16.05 | NS | 57 | 39 | 15.05 | 14.59 | NS | 100 | 26 | 14.55 | 14.10 | NS |
| (1) Superior Girls vs.<br>(2) Average Girls | 54 | 89 | 13.94 | 14.47 | NS | 59 | 59 | 15.46 | 13.81 | NS | 63 | 41 | 14.89 | 13.51 | NS |
| (1) Superior Girls vs.<br>(2) Slow Girls | 54 | 35 | 13.94 | 13.88 | NS | 59 | 33 | 15.46 | 15.03 | NS | 63 | 37 | 14.89 | 14.57 | NS |
| (1) Average Girls vs.<br>(2) Slow Girls | 89 | 35 | 14.47 | 13.88 | NS | 59 | 33 | 13.81 | 15.03 | NS | 41 | 37 | 13.51 | 14.57 | NS |

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

Table 12-10

Within Treatment Differences on the GAMIN Inferiority Score

| Group | Sample VI | | | | | Sample VII | | | | | Sample VIII | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $N_1$ | $N_2$ | $M_1$ | $M_2$ | t | $N_1$ | $N_2$ | $M_1$ | $M_2$ | t | $N_1$ | $N_2$ | $M_1$ | $M_2$ | t |
| District A | | | | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 65 | 172 | 24.51 | 22.59 | NS[b] | 49 | 138 | 26.16 | 23.14 | 2.02 | 47 | 123 | 25.85 | 22.85 | AS[a] |
| (1) Superior Boys vs.<br>(2) Slow Boys | 65 | 35 | 24.51 | 21.29 | AS | 49 | 36 | 26.16 | 21.28 | 2.61 | 47 | 28 | 25.85 | 26.57 | NS |
| (1) Average Boys vs.<br>(2) Slow Boys | 172 | 35 | 22.59 | 21.29 | NS | 138 | 36 | 23.14 | 21.28 | NS | 123 | 28 | 22.85 | 26.57 | AS |
| (1) Superior Girls vs.<br>(2) Average Girls | 58 | 92 | 25.29 | 22.96 | AS | 60 | 116 | 22.80 | 21.06 | NS | 85 | 31 | 23.31 | 22.37 | NS |
| (1) Superior Girls vs.<br>(2) Slow Girls | 58 | 58 | 25.29 | 22.62 | 2.72 | 60 | 18 | 22.80 | 17.50 | 2.29 | 85 | 40 | 23.31 | 22.65 | NS |
| (1) Average Girls vs.<br>(2) Slow Girls | 92 | 58 | 22.96 | 22.62 | AS | 116 | 18 | 21.06 | 17.50 | AS | 31 | 40 | 22.37 | 22.65 | NS |
| District R | | | | | | | | | | | | | | | |
| (1) Superior Boys vs.<br>(2) Average Boys | 81 | 175 | 24.19 | 22.09 | AS | 79 | 57 | 24.57 | 22.46 | NS | 62 | 100 | 25.02 | 22.71 | NS |
| (1) Superior Boys vs.<br>(2) Slow Boys | 81 | 23 | 24.19 | 18.70 | NS | 79 | 39 | 24.57 | 21.05 | NS | 62 | 26 | 25.02 | 21.08 | NS |
| (1) Average Boys vs.<br>(2) Slow Boys | 175 | 23 | 22.09 | 18.70 | NS | 57 | 39 | 22.46 | 21.05 | NS | 100 | 26 | 22.71 | 21.08 | NS |
| (1) Superior Girls vs.<br>(2) Average Girls | 54 | 89 | 24.80 | 22.76 | NS | 59 | 59 | 26.46 | 20.85 | 3.38 | 63 | 41 | 24.92 | 19.93 | 2.48 |
| (1) Superior Girls vs.<br>(2) Slow Girls | 54 | 35 | 24.80 | 20.69 | 2.28 | 59 | 33 | 26.46 | 22.70 | AS | 63 | 37 | 24.92 | 20.11 | 2.56 |
| (1) Average Girls vs.<br>(2) Slow Girls | 89 | 35 | 22.76 | 20.69 | NS | 59 | 33 | 20.85 | 22.70 | NS | 41 | 37 | 19.93 | 20.11 | NS |

---

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

and girls reaching or approaching significance. Slow boys and girls differed least on the Ascendancy score, with only one of the differences at this level reaching statistical significance. It appears safe to conclude that there is a consistent tendency for boys to obtain higher Ascendancy scores than girls during the junior high school grades.

### Inferiority

Between-treatment differences on the Inferiority factor of the GAMIN are summarized in Table 12-8. It was hypothesized that superior District A pupils would receive less favorable scores on this variable than comparable District R pupils. Although a slight trend was evident in the hypothesized direction, only one of the six comparisons between scores of superior pupils on this factor reached statistical significance. It was also hypothesized that slow pupils in District A would receive less favorable scores on this variable than their District R counterparts. The results showed no consistent pattern of differences favoring either grouping treatment for slow pupils on the Inferiority factor. Two of the six comparisons involving slow groups in the three samples were significant, one favoring District R slow girls in Sample VII and the other favoring District A slow boys in Sample VIII. Of the four remaining differences between slow pupils, two favored District R and two favored District A. These data would lead to the conclusion that feelings of inferiority as measured by GAMIN Factor I are not consistently different in ability grouped and random grouped classrooms.

Within-treatment differences in Factor I (Table 12-10) reflected a tendency for pupils of higher ability in both districts to obtain more favorable scores. Four of the 18 comparisons in District A and five of the 18 comparisons in District R were statistically significant.

Sex differences on this variable are summarized in Table 12-11. Of the 18 comparisons made, none reached statistical significance, thus leading to the conclusion that at the junior high school level there are no true differences between comparable girls' and boys' groups in inferiority feelings as measured by Factor I of the GAMIN.

## Objective-Analytic Personality Test Factors (OA)

Analysis of a data as obtained on the experimental form of the battery required considerable effort. Cattell (1955) recommends that the various tests going into a factor score be weighted equally, thus making it necessary to convert the raw scores on each test into standard scores and then add the standard scores for the subtests used in order to obtain a factor score. This procedure was followed to obtain a factor score for Factor UI-16, Competent Assertiveness. The performance of various Sample IV subgroups on this factor score was then compared using analysis of variance.

### Competent Assertiveness

In general, comparisons between similar pupils in the two grouping treatments revealed very similar means on this factor score. Only one of the six between-treatment comparisons revealed a statistically significant difference, District R slow girls obtaining higher mean score (t=3.90) than District A slow girls.

Differences between pupils of different ability levels generally indicated a

Table 12-11

Sex Differences on the GAMIN Ascendancy and Inferiority Scores

| Group | Sample VI |  |  | Sample VII |  |  | Sample VIII |  |  |
|---|---|---|---|---|---|---|---|---|---|
|  | Mean A | Mean R | t | Mean A | Mean R | t | Mean A | Mean R | t |
|  |  |  |  | **Ascendancy** |  |  |  |  |  |
| **District A** |  |  |  |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 16.05 | 13.81 | 2.50 | 16.35 | 14.32 | AS[a] | 17.23 | 13.92 | 2.87 |
| Average Boys vs. Average Girls | 15.82 | 13.65 | 3.11 | 15.71 | 14.28 | 2.12 | 15.99 | 14.18 | 2.21 |
| Slow Boys vs. Slow Girls | 16.03 | 10.09 | 2.35 | 14.83 | 13.17 | NS[b] | 15.43 | 13.12 | AS |
| **District R** |  |  |  |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 18.05 | 13.94 | 4.02 | 16.07 | 15.46 | NS | 17.54 | 14.89 | 2.37 |
| Average Boys vs. Average Girls | 16.35 | 14.47 | 2.34 | 15.05 | 13.81 | NS | 14.55 | 13.51 | NS |
| Slow Boys vs. Slow Girls | 16.05 | 13.89 | AS | 14.59 | 15.03 | NS | 14.10 | 14.57 | NS |
|  |  |  |  | **Inferiority** |  |  |  |  |  |
|  | Sample VI |  |  | Sample VII |  |  | Sample VIII |  |  |
| **District A** |  |  |  |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 24.51 | 24.10 | NS | 26.16 | 22.80 | AS | 25.85 | 23.31 | NS |
| Average Boys vs. Average Girls | 22.59 | 22.09 | NS | 23.14 | 21.06 | AS | 22.85 | 22.37 | NS |
| Slow Boys vs. Slow Girls | 21.29 | 18.70 | NS | 21.28 | 17.50 | AS | 26.57 | 22.65 | NS |
| **District R** |  |  |  |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 25.29 | 24.80 | NS | 24.57 | 26.46 | NS | 25.02 | 24.92 | NS |
| Average Boys vs. Average Girls | 22.96 | 22.76 | NS | 22.46 | 20.85 | NS | 22.71 | 19.93 | NS |
| Slow Boys vs. Slow Girls | 22.62 | 20.69 | NS | 21.05 | 22.70 | NS | 21.08 | 20.11 | NS |

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

positive relationship between Competent Assertiveness and ability level in both districts. Eleven of the 12 within-treatment differences were statistically significant, all favoring the higher of the two groups being compared.

Comparisons between boys and girls of comparable ability in the two districts showed girls to have higher scores on this factor in all comparisons with four of the six differences reaching statistical significance.

## Anxiety to Achieve

In order to gain some insights into the specific measures, the six tests selected to measure Cattell's factor UI-24, Anxiety to Achieve, were analysed individually instead of combining them into a single factor score. Between-treatment comparisons on these six tests for superior pupils revealed not a single statistically significant difference among the 12 comparisons made. Similarly, comparisons for slow pupils failed to reveal a significant difference. Only the comparisons between average boys in the two districts revealed a consistent pattern of significant difference. Of the six between-treatment comparisons made for this group, four reached statistical significance, all favoring the District R pupils.[1] One significant difference on Test G-47 was found in the between-treatment comparison of average girls, again favoring the District R sample.

Comparisons between pupils of different ability levels within each district revealed a tendency for more able pupils to obtain higher scores on the tests employed to measure Anxiety to Achieve. This difference was most marked for Test G-29 where all 12 within-treatment comparisons were statistically significant and Test G-47 where 11 of the 12 within-treatment comparisons reached significance. The differences were smallest for Tests G-41 and G-38 where only one of the 12 within-treatment comparisons made for each test reached statistical significance.

No consistent overall pattern of sex differences is evident in the six tests selected to measure the Anxiety to Achieve factor. On Test G-47, five of the six sex differences were statistically significant all favoring the girl's groups. No significant sex differences were present on Test G-31 and G-38; one difference favoring girls on Test G-29, and four differences all favoring boys on Test G-41. If a factor score had been obtained, the differences favoring girls on Test G-47 would probably have cancelled those favoring boys on Test G-41.

Interpretation of the results on the Cattell factors is difficult, as would be expected from the exploratory nature of the measures involved. The general lack of consistent between-treatment differences, however, suggests that whatever these tests measure is not being differentially affected by the ability grouping and random grouping treatments. As more experience is gained with these measures, more meaningful interpretations will be possible.

---

[1] Tests G-47, G-29, G-47 and G-41.

## California Psychological Inventory (CPI)

### Class I Variables

Between-treatment differences on the Class I variables of the CPI are summarized in Tables 12-12 and 12-13. The reader will recall that the six personality variables included in this class are broadly defined by the test author as measures of poise, ascendancy, and self-assurance. Of the 24 comparisons between superior pupils in the two districts summarized on these tables, three reached statistical significance all favoring Sample VII superior boys in District A over comparable District R pupils. As no significant differences occurred between superior pupils in Sample VI, the Sample VII results must be considered tentative.

Comparisons between slow pupils in the two districts revealed only two significant differences, one favoring District R pupils and one favoring District A pupils. This would suggest that neither treatment has any significant influence upon poise, ascendancy, and self-assurance of slow pupils.

Comparisons between average pupils revealed eight of the 24 between-treatment comparisons for average pupils to be statistically significant. All eight of these differences indicated higher scores for District R pupils. Five of these differences occurred in Sample VI and three in Sample VII, suggesting some consistency in the results. No significant differences on these variables were hypothesized for average pupils and the investigator has no explanation for the fairly consistent tendency for District R pupils of average ability to attain more favorable scores on the Class I CPI personality variables.

Within-treatment differences on the CPI variables are summarized in Tables 12-14 and 12-15. It will be noted that in comparisons between pupils of different ability levels there was a consistent tendency for pupils of higher ability to obtain higher scores on the Class I variables. Of the 36 comparisons between District A pupils of different ability levels in Sample VI, 23 were statistically significant, all favoring the group of higher ability. Closely comparable results were obtained for District A Sample VII where 26 of the 36 within-treatment comparisons were statistically significant, again favoring the group of higher ability. District R within-treatment comparisons followed the same trend although differences were generally smaller. For Sample VI pupils in District R, 15 of the within-treatment comparisons were statistically significant while 19 significant differences were obtained from the 36 comparisons between Sample VII subgroups. Thus, it may be concluded that ability level in both treatments is positively related to scores on the Class I personality variables of the CPI, with this relationship apparently somewhat more pronounced for pupils in ability grouped classes.

A total of 36 comparisons were made between boys and girls of comparable ability level in each district. For Sample VI pupils, there were six significant sex differences, five of them favoring girls. For Sample VII, girls obtained significantly higher scores in three comparisons; boys in two comparisons. Significant sex differences were generally scattered with no pattern emerging that would suggest true generalizable differences between the sexes on these variables (see Tables 12-20 and 12-21).

### Table 12-12

Sample VI Between-Treatment Differences on Class I Variables of the CPI

| | | | Dominance | | | Capacity for Status | | | Sociability | | | Social Presence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | $N_A$ | $N_R$ | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys | | | | | | | | | | | | | | |
| Dist. A vs. R | 72 | 116 | 25.81 | 24.50 | N.S.[b] | 15.12 | 15.03 | N.S. | 23.72 | 22.74 | N.S. | 32.89 | 32.99 | N.S. |
| Superior Girls | | | | | | | | | | | | | | |
| Dist. A vs. R | 90 | 88 | 26.23 | 24.98 | N.S. | 15.87 | 15.65 | N.S. | 23.72 | 23.05 | N.S. | 31.64 | 31.59 | N.S. |
| Average Boys | | | | | | | | | | | | | | |
| Dist. A vs. R | 137 | 65 | 23.36 | 24.02 | N.S. | 14.21 | 14.40 | N.S. | 21.68 | 22.82 | N.S. | 31.54 | 33.78 | 2.96 |
| Average Girls | | | | | | | | | | | | | | |
| Dist. A vs. R | 129 | 54 | 23.64 | 24.78 | N.S. | 14.72 | 16.15 | N.S. | 21.53 | 23.19 | A.S.[a] | 31.29 | 31.57 | N.S. |
| Slow Boys | | | | | | | | | | | | | | |
| Dist. A vs. R | 46 | 44 | 22.85 | 21.48 | N.S. | 14.04 | 13.54 | N.S. | 20.98 | 19.57 | N.S. | 30.41 | 30.43 | N.S. |
| Slow Girls | | | | | | | | | | | | | | |
| Dist. A vs. R | 44 | 43 | 22.66 | 23.79 | N.S. | 13.50 | 14.65 | N.S. | 19.84 | 20.77 | N.S. | 28.95 | 30.70 | A.S. |

| Group | $N_A$ | $N_R$ | Self Acceptance $M_A$ | $M_R$ | t | Sense of Well-Being $M_A$ | $M_R$ | t |
|---|---|---|---|---|---|---|---|---|
| Superior Boys<br>Dist. A vs. R | 72 | 116 | 18.97 | 19.21 | N.S. | 31.44 | 30.69 | N.S. |
| Superior Girls<br>Dist. A vs. R | 90 | 88 | 19.41 | 19.43 | N.S. | 34.33 | 32.64 | A.S. |
| Average Boys<br>Dist. A vs. R | 137 | 65 | 17.87 | 19.35 | 2.98 | 29.47 | 30.00 | N.S. |
| Average Girls<br>Dist. A vs. R | 129 | 54 | 17.84 | 19.33 | 2.74 | 30.18 | 32.48 | 2.01 |
| Slow Boys<br>Dist. A vs. R | 46 | 44 | 17.17 | 17.73 | N.S. | 26.33 | 25.95 | N.S. |
| Slow Girls<br>Dist. A vs. R | 44 | 43 | 16.70 | 18.21 | A.S. | 26.36 | 29.72 | 2.00 |

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

Table 12-13

Sample VII Between-Treatment Differences on Class I Variables of the CPI

| | | | Dominance | | | Capacity for Status | | | Sociability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | $N_A$ | $N_R$ | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys<br>Dist. A vs. R | 39 | 64 | 26.38 | 24.68 | N.S.[b] | 16.82 | 15.05 | 2.53 | 24.74 | 23.13 | N.S. |
| Superior Girls<br>Dist. A vs. R | 61 | 54 | 24.67 | 26.30 | N.S. | 15.72 | 17.37 | A.S.[a] | 22.41 | 23.68 | N.S. |
| Average Boys<br>Dist. A vs. R | 123 | 44 | 22.22 | 23.11 | N.S. | 13.71 | 15.23 | 2.42 | 21.06 | 22.16 | N.S. |
| Average Girls<br>Dist. A vs. R | 101 | 43 | 23.06 | 25.47 | 2.37 | 14.95 | 15.28 | N.S. | 20.99 | 22.30 | N.S. |
| Slow Boys<br>Dist. A vs. R | 39 | 30 | 20.69 | 21.27 | N.S. | 12.77 | 13.13 | N.S. | 18.03 | 20.23 | 2.33 |
| Slow Girls<br>Dist. A vs. R | 15 | 28 | 19.73 | 21.46 | N.S. | 12.87 | 13.11 | N.S. | 18.53 | 18.93 | N.S. |

| Group | $N_A$ | $N_R$ | Social Presence $M_A$ | Social Presence $M_R$ | Social Presence t | Self-Acceptance $M_A$ | Self-Acceptance $M_R$ | Self-Acceptance t | Sense of Well-Being $M_A$ | Sense of Well-Being $M_R$ | Sense of Well-Being t |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Superior Boys<br>Dist. A vs. R | 39 | 64 | 34.97 | 32.67 | 2.05 | 20.23 | 19.28 | N.S. | 34.38 | 31.23 | 2.59 |
| Superior Girls<br>Dist. A vs. R | 61 | 54 | 32.02 | 32.56 | N.S. | 19.08 | 20.26 | N.S. | 33.02 | 33.44 | N.S. |
| Average Boys<br>Dist. A vs. R | 123 | 44 | 31.89 | 31.70 | N.S. | 17.65 | 18.64 | N.S. | 29.22 | 27.70 | N.S. |
| Average Girls<br>Dist. A vs. R | 101 | 43 | 30.74 | 32.74 | A.S. | 17.83 | 19.79 | 2.84 | 30.75 | 30.21 | N.S. |
| Slow Boys<br>Dist. A vs. R | 39 | 30 | 29.46 | 30.83 | N.S. | 16.49 | 17.23 | N.S. | 24.79 | 25.57 | N.S. |
| Slow Girls<br>Dist. A vs. R | 15 | 28 | 26.93 | 30.14 | A.S. | 15.47 | 17.29 | N.S. | 25.67 | 26.96 | N.S. |

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

## Table 12-14

Significant Level Differences on the CPI Variables for Sample VI

| | Class I | | | | | | Class II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Do | Cs | Sy | Sp | Sa | Wb | Re | So | Sc | To | Gi | Cm |
| **District A** | | | | | | | | | | | | |
| (1) Superior Boys vs. | 1[a] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 1 |
| (2) Average Boys | 3.32 | A.S.[b] | 2.92 | A.S. | 2.21 | 2.01 | 4.51 | 3.79 | | | | 5.16 |
| (1) Superior Boys vs. | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Boys | 2.27 | | 2.96 | 2.71 | 2.81 | 3.92 | 4.51 | 3.38 | | 2.95 | | 4.27 |
| (1) Average Boys vs. | | | | | | 1 | | | 1 | 1 | 1 | |
| (2) Slow Boys | | | | | | 2.55 | | | A.S. | 2.02 | A.S. | |
| (1) Superior Girls vs. | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| (2) Average Girls | 3.46 | 2.03 | 3.07 | | 3.34 | 4.42 | 4.22 | 3.32 | 3.17 | 4.10 | 3.18 | 3.65 |
| (1) Superior Girls vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| (2) Slow Girls | 3.21 | 3.21 | 4.19 | 2.67 | 3.80 | 6.27 | 6.06 | 5.74 | 3.82 | 5.14 | 2.97 | 6.51 |
| (1) Average Girls vs. | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 |
| (2) Slow Girls | | A.S. | A.S. | 2.66 | A.S. | 2.91 | 2.64 | 4.18 | A.S. | 2.57 | | 3.80 |
| **District R** | | | | | | | | | | | | |
| (1) Superior Boys vs. | | | | | | | 1 | | | | | 1 |
| (2) Average Boys | | | | | | | 1.96 | | | | | 2.17 |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Boys | 3.64 | 2.38 | 3.59 | 2.55 | 1.96 | 4.33 | 6.13 | 4.12 | | 4.60 | | 5.98 |
| (1) Average Boys vs. | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Boys | 4.45 | | 3.65 | 3.14 | 2.50 | 3.10 | 3.94 | 2.44 | | 3.16 | | 4.00 |
| (1) Superior Girls vs. | | | | | | | 1 | | | | | |
| (2) Average Girls | | | | | | | 2.68 | | | | | |
| (1) Superior Girls vs. | | | 1 | | | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Girls | | | 2.52 | | | 2.22 | 5.13 | 2.14 | | 2.78 | | 2.92 |
| (1) Average Girls vs. | | 1 | 1 | | | 1 | 1 | 1 | | 1 | | |
| (2) Slow Girls | | 1.97 | 2.42 | | | A.S. | 2.41 | A.S. | | A.S. | | |

|  | Class III |  |  | Class IV |  |  |
|---|---|---|---|---|---|---|
| Group | Ac | Ai | Ie | Py | Fx | Fe |
| **District A** |  |  |  |  |  |  |
| (1) Superior Boys vs. | 1 | 1 | 1 |  |  |  |
| (2) Average Boys | 2.18 | A.S. | 3.66 |  |  |  |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 |  |  |
| (2) Slow Boys | 3.41 | 2.03 | 3.91 | A.S. |  |  |
| (1) Average Boys vs. | 1 |  |  | 1 |  | 2 |
| (2) Slow Boys | 2.33 |  |  | A.S. |  | A.S. |
| (1) Superior Girls vs. | 1 | 1 | 1 | 1 |  |  |
| (2) Average Girls | 4.49 | 3.56 | 4.95 | 2.06 |  |  |
| (1) Superior Girls vs. | 1 | 1 | 1 | 1 |  | 1 |
| (2) Slow Girls | 5.85 | 3.16 | 7.15 | 2.52 |  | 2.62 |
| (1) Average Girls vs. | 1 |  | 1 |  |  | 1 |
| (2) Slow Girls | 2.46 |  | 3.78 |  |  | 2.48 |
| **District R** |  |  |  |  |  |  |
| (1) Superior Boys vs. |  |  | 1 |  |  |  |
| (2) Average Boys |  |  | 2.86 |  |  |  |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 |  |  |
| (2) Slow Boys | 4.12 | 2.83 | 6.23 | 2.38 |  |  |
| (1) Average Boys vs. | 1 | 1 | 1 | 1 |  |  |
| (2) Slow Boys | 3.53 | A.S. | 3.11 | 2.38 |  |  |
| (1) Superior Girls vs. |  |  |  |  |  |  |
| (2) Average Girls |  |  |  |  |  |  |
| (1) Superior Girls vs. | 1 | 1 | 1 |  |  | 1 |
| (2) Slow Girls | 3.63 | 2.23 | 4.63 |  |  | 3.01 |
| (1) Average Girls vs. | 1 |  | 1 |  |  | 1 |
| (2) Slow Girls | 2.13 |  | 3.57 |  |  | 2.58 |

---

[a] Indicates that of the two groups compared, group 1, in this case Superior Boys obtained a significantly higher score on Dominance (Do) than group 2 (Average Boys).

[b] A.S. indicates difference approached significance, t=1.65-1.95.

## Table 12-15

### Significant Level Differences on the CPI Variables for Sample VII

| | Class I | | | | | Class II | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Do | Cs | Sy | Sp | Sa | Wb | Re | So | Sc | To | Gi | Cm |
| **District A** | | | | | | | | | | | | |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Average Boys | 3.79 | 4.58 | 3.77 | 2.88 | 3.52 | 3.48 | 5.66 | 5.97 | | 4.68 | | 4.88 |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Boys | 4.23 | 4.75 | 6.17 | 4.65 | 4.27 | 6.19 | 7.89 | 6.89 | | 5.57 | | 8.05 |
| (1) Average Boys vs. | | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Boys | | | 3.26 | 2.37 | A.S.[a] | 2.78 | 3.96 | 2.04 | | 2.56 | | 3.42 |
| (1) Superior Girls vs. | 1 | | 1 | | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Average Girls | A.S. | | A.S. | | A.S. | 2.09 | 3.98 | 2.27 | | 1.98 | | 4.03 |
| (1) Superior Girls vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Girls | 2.90 | 2.19 | 2.63 | 3.09 | 2.99 | 3.58 | 6.71 | 4.20 | | 5.18 | | 7.94 |
| (1) Average Girls vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Girls | 2.15 | A.S. | A.S. | 2.27 | 2.13 | 2.72 | 3.18 | 2.91 | | 2.53 | | 2.68 |
| **District R** | | | | | | | | | | | | |
| (1) Superior Boys vs. | | | | | | 1 | 1 | 1 | 1 | | | 1 |
| (2) Average Boys | | | | | | 2.50 | 3.15 | 2.20 | A.S. | | | 3.96 |
| (1) Superior Boys vs. | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Boys | 2.91 | 2.85 | 2.97 | | 2.28 | 3.76 | 6.37 | 5.19 | | 3.24 | | 6.41 |
| (1) Average Boys vs. | | 1 | 1 | | 1 | | 1 | 1 | 1 | 1 | | |
| (2) Slow Boys | | 2.76 | A.S. | | A.S. | | 2.96 | 2.57 | A.S. | 2.01 | | |
| (1) Superior Girls vs. | | 1 | | | | 1 | | | 1 | 1 | | |
| (2) Average Girls | | 2.40 | | | | 2.37 | | | A.S. | 2.01 | | |
| (1) Superior Girls vs. | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | | 1 | | 1 |
| (2) Slow Girls | 3.52 | 4.36 | 4.07 | A.S. | | 3.64 | 3.99 | 1.96 | | 3.51 | | A.S. |
| (1) Average Girls vs. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 1 | | 1 |
| (2) Slow Girls | 3.11 | 2.08 | 2.84 | 2.22 | 2.74 | A.S. | 3.94 | | | 2.14 | | 3.20 |

| Group | Class III<br>Ac | Ai | Ie | Class IV<br>Py | Fx | Fe |
|---|---|---|---|---|---|---|
| District A | | | | | | |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 | | |
| (2) Average Boys | 5.60 | 4.53 | 6.28 | A.S. | | |
| (1) Superior Boys vs. | 1 | 1 | 1 | 1 | | |
| (2) Slow Boys | 6.32 | 3.83 | 8.51 | A.S. | | |
| (1) Average Boys vs. | 1 | | 1 | | | 2 |
| (2) Slow Boys | A.S. | | 3.27 | | | A.S. |
| (1) Superior Girls vs. | 1 | 1 | 1 | | | |
| (2) Average Girls | 2.31 | 2.42 | 3.61 | | | |
| (1) Superior Girls vs. | 1 | 1 | 1 | | 1 | |
| (2) Slow Girls | 2.69 | 3.13 | 6.00 | | 2.23 | |
| (1) Average Girls vs. | | | 1 | | 1 | |
| (2) Slow Girls | | | 3.54 | | A.S. | |
| District R | | | | | | |
| (1) Superior Boys vs. | 1 | | 1 | | | |
| (2) Average Boys | A.S. | | 4.10 | | | |
| (1) Superior Boys vs. | 1 | 1 | 1 | | | |
| (2) Slow Boys | 4.40 | 2.21 | 5.51 | | | |
| (1) Average Boys vs. | 1 | | | | | |
| (2) Slow Boys | 2.65 | | | | | |
| (1) Superior Girls vs. | | | | | 2 | |
| (2) Average Girls | | | | | 2.71 | |
| (1) Superior Girls vs. | 1 | 1 | 1 | | | |
| (2) Slow Girls | 2.86 | 3.11 | 3.44 | | | |
| (1) Average Girls vs. | 1 | 1 | 1 | 1 | | |
| (2) Slow Girls | A.S. | 2.67 | 3.54 | A.S. | | |

[a] A.S. indicates difference approached significance, t=1.65-1.95.

## Class II Variables

The six Class II measures on the CPI fall into the broad area of socialization, maturity, and responsibility. Between-treatment differences on these measures are summarized in Tables 12-16 and 12-17. Comparisons between superior pupils in the two treatments showed two significant differences in each sample favoring District A. These four differences out of a total of 24 comparisons suggest a slight tendency for District A superior pupils to obtain more favorable scores in this broad personality area. Only two significant differences were found between average pupils in the two districts, one favoring District R and one favoring District A. Comparisons between slow pupils also failed to reveal any consistent trend on the Class II variables with two significant differences favoring District A and one favoring District R.

Within-treatment differences on the six Class II variables of the CPI are summarized in Tables 12-14 and 12-15. On four of these six variables; Responsibility, Socialization, Tolerance, and Communality, a very marked tendency for pupils of higher ability levels to obtain higher scores is in evidence. Comparisons between District A pupils in Sample VI revealed 20 of the 24 differences between pupils of different ability levels on these four variables to be statistically significant. Similar results were obtained for District A pupils in Sample VII, where all 24 of the within-treatment differences on the aforementioned variables were statistically significant. Two of the Class II variables, Self-Control and Good Impression, showed no consistent relationship with ability level. Within-treatment comparisons on these variables for Sample VI revealed only four significant differences, all in District A, while no significant differences were found on these variables for Sample VII.

District R within-treatment comparisons generally followed the same pattern found in District A except that differences were less pronounced between different ability levels. In Sample VI, 17 of these differences were statistically significant, all favoring the group of higher ability. There were also 17 significant within-treatment differences in the Sample VII data for District R. Within-treatment results on the Class II variables were thus quite similar to those obtained on the Class I variables, with a consistent relationship being found in both districts between ability level and favorable scores. For both Class I and Class II variables this relationship was more pronounced in District A.

Sex differences on the Class II CPI variables generally favored girls in both districts and both samples. For District A pupils, nine of the 18 sex differences in Sample VI were statistically significant. All of the differences between superior pupils were statistically significant and three of the six differences between average pupils reached significance. For Sample VII there were five significant differences between District A boys and girls, all occurring between boys and girls of average ability.

Sex differences in District R were somewhat more pronounced with 14 of the 18 differences in Sample VI being statistically significant. In Sample VII, eight of the 18 differences between comparable pupils of the opposite sex reached statistical significance.

In summary, girls were consistently higher on the Class II personality variables in both treatments, with differences between the sexes being somewhat more pronounced in District R.

Table 12-16

Sample VI Between-Treatment Differences on the Class II Variables of the CPI

| Group | $N_A$ | $N_R$ | Responsibility $M_A$ | Responsibility $M_R$ | Responsibility t | Socialization $M_A$ | Socialization $M_R$ | Socialization t | Self Control $M_A$ | Self Control $M_R$ | Self Control t |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Superior Boys Dist. A vs. R | 72 | 116 | 28.08 | 26.53 | A.S.[a] | 37.54 | 35.83 | A.S. | 24.49 | 23.96 | N.S.[b] |
| Superior Girls Dist. A vs. R | 90 | 88 | 30.51 | 30.88 | N.S. | 40.82 | 40.11 | N.S. | 30.18 | 27.39 | 2.21 |
| Average Boys Dist. A vs. R | 137 | 65 | 24.58 | 24.93 | N.S. | 33.80 | 34.28 | N.S. | 24.96 | 23.80 | N.S. |
| Average Girls Dist. A vs. R | 129 | 54 | 27.36 | 28.76 | N.S. | 38.21 | 39.57 | N.S. | 26.64 | 27.15 | N.S. |
| Slow Boys Dist. A vs. R | 46 | 44 | 23.61 | 21.30 | 2.71 | 33.28 | 31.52 | N.S. | 22.60 | 23.00 | N.S. |
| Slow Girls Dist. A vs. R | 44 | 43 | 24.93 | 26.21 | N.S. | 33.86 | 37.37 | 2.26 | 24.30 | 27.15 | N.S. |

| Group | $N_A$ | $N_R$ | Tolerance $M_A$ | Tolerance $M_R$ | Tolerance t | Good Impression $M_A$ | Good Impression $M_R$ | Good Impression t | Communality $M_A$ | Communality $M_R$ | Communality t |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Superior Boys<br>Dist. A vs. R | 72 | 116 | 17.50 | 17.27 | N.S. | 15.78 | 14.57 | N.S. | 24.93 | 23.84 | A.S. |
| Superior Girls<br>Dist. A vs. R | 90 | 88 | 19.89 | 18.38 | A.S. | 18.72 | 16.25 | 2.55 | 25.54 | 25.41 | N.S. |
| Average Boys<br>Dist. A vs. R | 137 | 65 | 16.27 | 16.42 | N.S. | 16.72 | 15.88 | N.S. | 21.64 | 22.38 | N.S. |
| Average Girls<br>Dist. A vs. R | 129 | 54 | 16.91 | 17.74 | N.S. | 15.81 | 17.39 | N.S. | 23.99 | 24.41 | N.S. |
| Slow Boys<br>Dist. A vs. R | 46 | 44 | 14.48 | 13.50 | N.S. | 14.96 | 15.30 | N.S. | 21.34 | 19.02 | 2.25 |
| Slow Girls<br>Dist. A vs. R | 44 | 43 | 14.64 | 15.63 | N.S. | 15.25 | 17.19 | N.S. | 21.25 | 23.19 | A.S. |

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

Table 12-17

Sample VII Between Treatment Differences on Class II Variables of the CPI

| Group | Responsibility | | | Socialization | | | Self Control | | |
|---|---|---|---|---|---|---|---|---|---|
| | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys<br>Dist. A vs. R | 27.74 | 28.03 | A.S.[a] | 40.00 | 36.72 | 2.50 | 29.79 | 34.36 | N.S.[b] |
| Superior Girls<br>Dist. A vs. R | 31.43 | 30.72 | N.S. | 39.69 | 39.33 | N.S. | 37.87 | 35.11 | N.S. |
| Average Boys<br>Dist. A vs. R | 24.33 | 24.77 | N.S. | 32.61 | 33.43 | N.S. | 25.76 | 26.16 | N.S. |
| Average Girls<br>Dist. A vs. R | 28.37 | 29.02 | N.S. | 37.26 | 37.56 | N.S. | 37.91 | 44.19 | N.S. |
| Slow Boys<br>Dist. A vs. R | 21.87 | 21.10 | N.S. | 29.87 | 28.87 | N.S. | 25.83 | 29.27 | N.S. |
| Slow Girls<br>Dist. A vs. R | 24.00 | 24.93 | N.S. | 31.27 | 35.75 | A.S. | 36.93 | 41.61 | N.S. |

| Group | Tolerance | | | Good Impression | | | Communality | | |
|---|---|---|---|---|---|---|---|---|---|
| | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys Dist. A vs. R | 20.28 | 17.56 | 2.83 | 15.00 | 14.77 | N.S. | 25.77 | 25.00 | N.S. |
| Superior Girls Dist. A vs. R | 20.44 | 19.39 | N.S. | 15.79 | 14.79 | N.S. | 26.07 | 25.63 | N.S. |
| Average Boys Dist. A vs. R | 16.09 | 17.98 | N.S. | 15.23 | 15.82 | N.S. | 21.32 | 20.71 | N.S. |
| Average Girls Dist. A vs. R | 18.14 | 16.88 | N.S. | 15.97 | 13.47 | 2.05 | 23.31 | 25.77 | 2.95 |
| Slow Boys Dist. A vs. R | 13.77 | 14.33 | N.S. | 15.38 | 15.70 | N.S. | 17.77 | 17.53 | N.S. |
| Slow Girls Dist. A vs. R | 12.53 | 14.18 | N.S. | 17.00 | 14.96 | N.S. | 19.40 | 23.18 | N.S. |

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

## Class III Variables

Class III personality variables on the CPI include measures of achievement potential and intellectual efficiency. Three scores are obtained in this area: Achievement via Conformance (Ac), Achievement via Independence (Ai), Intellectual Efficiency (Ie). Between-treatment differences on the Class III variables are summarized in Table 12-18. It was hypothesized that both superior and slow pupils would make higher scores on Achievement via Conformance in District A. One of the four comparisons between superior pupils on this variable reached statistical significance in the hypothesized direction. No significant differences between slow pupils were found. It was hypothesized that District A superior pupils would receive higher scores on Intellectual Efficiency. Although District A superior pupils obtained higher scores on all four of the comparisons on this variable, only one of the differences approached significance.

District R average girls in Sample VI obtained a significantly higher mean score on Intellectual Efficiency than comparable District A girls. It was also hypothesized that superior District R pupils would obtain more favorable scores on Achievement via Independence. The results, however, showed that District A superior pupils made higher scores on this variable in all four of the comparisons, one of which reached statistical significance. The only between-treatment results that were reasonably consistent on the Class III variable were between superior boys in Sample VII. District A pupils in this sample were significantly higher in Ac and Ai and also obtained higher scores, approaching significance on Ie.

As was hypothesized, within-treatment comparisons revealed a marked and consistent tendency for pupils of higher ability to obtain more favorable scores on all of the Class III variables. For District A, 29 of the 36 differences between pupils of different ability levels reported in Tables 12-14 and 12-15 were statistically significant, all favoring the groups of higher ability. A similar though less pronounced trend was found in the District R within-treatment comparisons, with 21 of the 36 differences being statistically significant, again favoring the group of higher ability. Thus, it appears that the Class III variables of the CPI are consistently related to ability level with this relationship being somewhat more pronounced in District A than in District R.

Sex difference comparisons on the Class III variables revealed a consistent pattern in both districts for girls to attain higher scores than boys. At the average ability level in District A, girls were significantly higher in all six comparisons while they were higher in four of the six comparisons in District R. Superior girls were significantly higher in three comparisons for District A and four comparisons for District R. Sex differences were considerably less pronounced among slow pupils with no significant differences occurring in District A and only three occurring in District R at this level.

In summary, all 20 of the significant differences between the sexes on the Class III variables favored girls, thus providing strong evidence that girls in both districts, especially at the superior and average ability levels, were higher on measures of achievement potential and intellectual efficiency.

## Class IV Variables

Class IV of the CPI contains three scores classified by the test author as measures of intellectual and interest modes. The three variables included in this class

Table 12-18

Between Treatment Differences on the Class III Variables of the CPI

| | Sample VI | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Achievement via Conformance | | | Achievement via Independence | | | Intellectual Efficiency | | |
| Group | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys<br>Dist. A vs. R | 23.32 | 22.81 | N.S.[b] | 14.63 | 14.41 | N.S. | 33.26 | 32.65 | N.S. |
| Superior Girls<br>Dist. A vs. R | 26.63 | 25.84 | N.S. | 16.51 | 15.86 | N.S. | 35.74 | 34.80 | N.S. |
| Average Boys<br>Dist. A vs. R | 21.69 | 22.35 | N.S. | 13.51 | 13.94 | N.S. | 29.98 | 29.83 | N.S. |
| Average Girls<br>Dist. A vs. R | 23.48 | 25.00 | A.S.[a] | 14.64 | 15.28 | N.S. | 31.60 | 34.09 | 2.61 |
| Slow Boys<br>Dist. A vs. R | 19.72 | 19.27 | N.S. | 12.89 | 12.57 | N.S. | 28.45 | 26.27 | A.S. |
| Slow Girls<br>Dist. A vs. R | 21.27 | 22.58 | N.S. | 13.61 | 13.98 | N.S. | 27.55 | 29.49 | N.S. |

| Group | Sample VII<br>Achievement via Conformance<br>$M_A$ | <br><br>$M_R$ | <br><br>t | <br>Achievement via Independence<br>$M_A$ | <br><br>$M_R$ | <br><br>t | <br>Intellectual Efficiency<br>$M_A$ | <br><br>$M_R$ | <br><br>t |
|---|---|---|---|---|---|---|---|---|---|
| Superior Boys<br>Dist. A vs. R | 25.81 | 23.27 | 2.49 | 16.53 | 14.63 | 2.52 | 36.41 | 34.09 | A.S. |
| Superior Girls<br>Dist. A vs. R | 24.70 | 25.52 | N.S. | 16.74 | 15.87 | N.S. | 35.89 | 35.04 | N.S. |
| Average Boys<br>Dist. A vs. R | 20.94 | 21.50 | N.S. | 13.60 | 13.98 | N.S. | 29.24 | 27.91 | N.S. |
| Average Girls<br>Dist. A vs. R | 22.72 | 23.72 | N.S. | 15.01 | 14.77 | N.S. | 31.96 | 32.91 | N.S. |
| Slow Boys<br>Dist. A vs. R | 19.36 | 18.30 | N.S. | 13.00 | 12.77 | N.S. | 25.41 | 26.30 | N.S. |
| Slow Girls<br>Dist. A vs. R | 20.67 | 21.50 | N.S. | 13.07 | 12.25 | N.S. | 24.87 | 28.25 | N.S. |

---

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

are: Psychological-mindedness (Py), Flexibility (Fx), and Femininity (Fe). It was hypothesized that District R pupils at all ability levels would develop higher scores on flexibility than comparable District A pupils. Table 12-19 summarizes between-treatment differences on the Class IV variables. It will be noted that two significant differences were found between superior pupils on Flexibility. As one of these significant differences favors each treatment, they tend to cancel each other out thus leading to rejection of the working hypothesis. No significant between-treatment differences emerged on the flexibility score for average and slow pupils.

Only one significant between-treatment differences emerged from the 12 comparisons on Psychological-mindedness. District A superior boys in Sample VII obtained a significantly higher score than comparable District R boys on this variable. The Sample VI data, however, show no such trend, suggesting that this was not a generalizable difference attributable to the treatments. No significant between-treatment differences were found in the 12 comparisons on the Femininity score of the CPI.

Within-treatment differences were much smaller on the Class IV variables of the CPI than those found for Classes I, II, III. In Sample VI, only four significant differences were found for each district on these three variables. All of these differences favored the group of higher ability. In Sample VII, only one significant difference was found in the District A comparisons and one in the District R comparisons. These data suggest that there is no consistent relationship between scores on the Class IV variables of the CPI and ability level.

Comparisons between the sexes on these variables revealed boys to be generally higher on Psychological-mindedness. Five of the 12 between-sex comparisons on this score were statistically significant, all favoring boys' groups. Only one significant sex difference was obtained from the 12 comparisons on Flexibility, although it may be seen in Tables 12-20 and 12-21 that a slight trend among the non-significant differences appears to favor the girls' groups. Sex difference scores on Femininity tend to provide some concurrent validity evidence for this scale of the CPI.

## SUMMARY AND CONCLUSIONS

### California Test of Personality (CTP)

This measure was administered to Sample IV only and results may be accepted with less confidence than those obtained from measures that were administered to two or more samples such as the GAMIN factors and the California Psychological Inventory.

It was hypothesized that superior and slow pupils in ability grouped classes would obtain higher scores on the Feeling of Belonging subscales. Results showed District R subgroups to be consistently higher on this score at all ability levels. These data suggest that ability grouping does not lead to a greater feeling of belonging on the part of pupils at any ability level but instead provides a less favorable climate than random grouping. Within-treatment comparisons on this variable showed a consistent trend in both districts for pupils of higher ability to obtain more favorable scores. It may be concluded that the pupil's feeling of belonging as measured by the CTP is more closely related to ability level than to the type of grouping treatment.

A lack of significant between-treatment differences on the Withdrawing Tendencies subscale suggests that this characteristic was not influenced by the grouping

## Table 12-19

Between Treatment Differences on Class IV Variables of the *CPI*

| | Sample VI | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Psychological-Mindedness | | | Flexibility | | | Femininity | | |
| Group | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys<br>Dist. A vs. R | 10.14 | 10.28 | N.S.[b] | 8.03 | 9.44 | 2.38 | 16.22 | 16.26 | N.S. |
| Superior Girls<br>Dist. A vs. R | 9.97 | 9.42 | N.S. | 9.08 | 9.30 | N.S. | 24.23 | 24.59 | N.S. |
| Average Boys<br>Dist. A vs. R | 10.05 | 10.13 | N.S. | 8.91 | 8.72 | N.S. | 16.22 | 16.03 | N.S. |
| Average Girls<br>Dist. A vs. R | 9.24 | 9.38 | N.S. | 9.63 | 9.30 | N.S. | 23.63 | 24.57 | A.S.[a] |
| Slow Boys<br>Dist. A vs. R | 9.22 | 8.93 | N.S. | 8.43 | 9.39 | N.S. | 17.33 | 16.14 | N.S. |
| Slow Girls<br>Dist. A vs. R | 8.77 | 9.09 | N.S. | 9.00 | 9.05 | N.S. | 22.07 | 22.44 | N.S. |

| | Sample VII | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Psychological-Mindedness | | | Flexibility | | | Femininity | | |
| Group | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Superior Boys<br>Dist. A vs. R | 10.82 | 9.67 | 2.41 | 9.33 | 8.88 | N.S. | 16.00 | 16.28 | N.S. |
| Superior Girls<br>Dist. A vs. R | 9.16 | 9.09 | N.S. | 9.80 | 8.06 | 2.61 | 24.03 | 22.56 | A.S. |
| Average Boys<br>Dist. A vs. R | 9.98 | 10.25 | N.S. | 8.31 | 8.05 | N.S. | 15.50 | 15.30 | N.S. |
| Average Girls<br>Dist. A vs. R | 8.92 | 9.30 | N.S. | 9.14 | 10.00 | N.S. | 23.84 | 23.67 | N.S. |
| Slow Boys<br>Dist. A vs. R | 9.67 | 10.20 | N.S. | 8.95 | 8.97 | N.S. | 16.56 | 16.50 | N.S. |
| Slow Girls<br>Dist. A vs. R | 8.73 | 8.07 | N.S. | 7.53 | 8.68 | N.S. | 24.13 | 23.39 | N.S. |

[a] A.S. indicates difference approached significance, t=1.65-1.95.
[b] N.S. indicates difference not significant.

Table 12-20

Sex Differences for Sample VI on the California Psychological Inventory

| | Class I | | | | | | Class II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Do | Cs | Sy | Sp | Sa | Wb | Re | So | Sc | To | Gi | Cm |
| District A | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | N.S.[b] | N.S. | N.S. | N.S. | N.S. | 2.96* | 2.77* | 3.25* | 4.17* | 2.70* | 2.92* | N.S. |
| Average Boys vs. Average Girls | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | 4.24* | 5.80* | A.S.*[a] | N.S. | N.S. | 4.65* |
| Slow Boys vs. Slow Girls | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. |
| District R | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | N.S. | N.S. | N.S. | N.S. | N.S. | 2.12* | 6.26* | 4.62* | 2.96* | N.S. | 2.00* | 2.55* |
| Average Boys vs. Average Girls | N.S. | N.S. | N.S. | 2.21 | N.S. | 1.96* | 4.08* | 4.66* | 2.26* | N.S. | N.S. | 2.87* |
| Slow Boys vs. Slow Girls | 2.35* | N.S. | N.S. | N.S. | N.S. | 2.60* | 4.82* | 4.65* | 2.12* | 1.98* | N.S. | 4.49* |

|  | Class III |  |  | Class IV |  |  |
|---|---|---|---|---|---|---|
| Group | Ac | Ai | Ie | Py | Fx | Fe |
| **District A** |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 3.95* | 2.71* | 2.49* | N.S. | A.S.* | 11.65* |
| Average Boys vs. Average Girls | 2.91* | 2.28* | 2.21* | 2.44 | N.S. | 16.98* |
| Slow Boys vs. Slow Girls | N.S. | N.S. | N.S. | N.S. | N.S. | 5.53* |
| **District R** |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 4.51* | 2.67* | 2.55* | 2.14 | N.S. | 17.43* |
| Average Boys vs. Average Girls | 2.87* | A.S.* | 3.77* | N.S. | N.S. | 12.65* |
| Slow Boys vs. Slow Girls | 3.06* | N.S. | 2.50* | N.S. | N.S. | 6.75* |

* Indicates girls made higher score.

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

7

Table 12-21

Sex Differences for Sample VII on the California Psychological Inventory

| Group | Class I | | | | | | Class II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Do | Cs | Sy | Sp | Sa | Wb | Re | So | Sc | To | Gi | Cm |
| **District A** | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | N.S.[b] | N.S. | 2.15 | 2.51 | N.S. | N.S. | A.S.* | N.S. | N.S. | N.S. | N.S. | N.S. |
| Average Boys vs. Average Girls | N.S. | 2.30* | N.S. | N.S. | N.S. | N.S. | 5.86* | 4.88* | 3.78* | 2.33* | N.S. | 2.74* |
| Slow Boys vs. Slow Girls | N.S. | N.S. | N.S. | A.S.[a] | N.S. | N.S. | A.S.* | N.S. | N.S. | N.S. | N.S. | N.S. |
| **District R** | | | | | | | | | | | | |
| Superior Boys vs. Superior Girls | N.S. | 3.44* | N.S. | N.S. | N.S. | A.S.* | 2.47* | A.S.* | N.S. | A.S.* | N.S. | N.S. |
| Average Boys vs. Average Girls | 2.97* | N.S. | N.S. | N.S. | N.S. | N.S. | 4.14* | 2.63* | 3.31* | N.S. | N.S. | 5.00* |
| Slow Boys vs. Slow Girls | N.S. | N.S. | N.S. | N.S. | N.S. | N.S. | 3.04* | 4.11* | A.S.* | N.S. | N.S. | 4.31* |

|  |  | Class III |  |  | Class IV |  |
|---|---|---|---|---|---|---|
| Group | Ac | Ai | Ie | Py | Fx | Fe |
| **District A** |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | N.S. | N.S. | N.S. | 3.04 | N.S. | 12.06* |
| Average Boys vs. Average Girls | 2.66* | 2.60* | 2.99* | 2.77 | A.S.* | 15.72* |
| Slow Boys vs. Slow Girls | N.S. | N.S. | N.S. | N.S. | N.S. | 7.82* |
| **District R** |  |  |  |  |  |  |
| Superior Boys vs. Superior Girls | 2.13* | N.S. | N.S. | N.S. | N.S. | 7.24* |
| Average Boys vs. Average Girls | 1.99* | N.S. | 3.12* | N.S. | 2.24* | 8.77* |
| Slow Boys vs. Slow Girls | 2.39* | N.S. | N.S. | 3.35 | N.S. | 8.16* |

* Indicates girls made higher score.

[a] A.S. indicates difference approached significance, t=1.65-1.95.

[b] N.S. indicates difference not significant.

treatments. Within-treatment differences showed a strong, consistent tendency for pupils of high ability to obtain more favorable scores leading to the conclusion that Withdrawing Tendencies as measured by the CTP were related to ability level but were not influenced by the grouping treatments.

Ability grouped pupils received less favorable scores on Anti-social Tendencies in all six between-treatment comparisons, with four differences reaching statistical significance. The results suggest that ability grouped pupils, particularly boys, report significantly more Anti-social Tendencies than do comparable pupils in random grouped classrooms. These differences were present at all three ability levels.

Total adjustment for the CTP subscales employed in the Utah study showed a clear tendency for random grouped pupils to obtain more favorable scores than ability grouped pupils. This tendency was significant in five of the six between-treatment comparisons and leads to the conclusion that random grouping provides a more favorable climate for total adjustment of the pupil as measured by the CTP than does ability grouping.

## Inventory of Factors GAMIN

Two personality factors from this inventory, Ascendancy and Inferiority, were measured for pupils in Samples VI, VII and VIII during the second year of the Utah study. These data, based on three samples, may be accepted with more confidence than the CTP data discussed above.

Of the 18 between-treatment comparisons made on the Ascendancy factor, only two reached statistical significance. These results suggest that the two treatments had no differential affect upon this personality variable. There was a slight but generally non-significant tendency in both districts for pupils of high ability to obtain higher scores on this variable. Between-sex comparisons showed a consistent trend in all three samples for boys to obtain more favorable Ascendancy scores than girls with nine of the 18 sex comparisons being statistically significant. We may conclude that boys at all ability levels are consistently more ascendant than girls during the junior high school grades. Sex differences were most pronounced between superior pupils, and with slow pupils differing least.

No consistent between-treatment differences emerged on the Inferiority factor of the GAMIN. These results support the conclusion that ability grouping is no more likely to develop inferiority feelings in pupils at any ability level than is random grouping. These data are particularly noteworthy in the case of slow pupils because of the frequently stated opinion that ability grouping does develop inferiority feelings among such pupils. These data, based on three different samples at different grade levels are more extensive than any previous evidence published concerning this question. The Utah study presents further evidence on this question in Chapter 13.

Within-treatment differences on the GAMIN Inferiority factor reflected a moderate tendency for pupils of higher ability in both districts to obtain more favorable scores. We may conclude that feelings of inferiority as measured by the GAMIN are inversely related to ability level regardless of the grouping treatment.

## Objective - Analytic Personality Test Factors (OA)

Two factors from this battery, Competent Assertiveness and Anxiety to Achieve, were administered in the Utah study; the former to Sample IV and the latter to Sample VI. As each factor was measured for only one sample, these results must be accepted with less confidence than results on the CPI and GAMIN for which data were available for two or more samples.

No consistent pattern of between-treatment differences emerged in either of these variables. Within-treatment differences on both factors were generally significant indicating a positive relationship between favorable scores on these factors and ability. Sex comparisons revealed that girls generally received higher scores on Competent Assertiveness than boys. From the above results we may conclude that Competent Assertiveness and Anxiety to Achieve as measured on Cattell's OA battery are not affected differentially by the ability grouping and random grouping treatments. We may also conclude that favorable scores on both factors are positively related to ability level.

California Psychological Inventory (CPI)

The CPI was administered to pupils in Sample VI and VII. The 18 CPI variables are classified into four broad areas. Class I variables include measures of poise, ascendancy, and self-assurance. Between-treatment comparisons on the Class I variables showed no consistent superiority for either treatment for superior pupils and slow pupils. Comparisons between average pupils revealed eight of the 24 between-treatment differences to be statistically significant, all favoring District R pupils. From these data we may conclude that the two grouping treatments do not appear to differentially affect personality variables related to poise, ascendancy, and self-assurance, excepting at the average ability level where random grouped pupils showed some tendency towards more favorable scores on these variables.

Within-treatment differences on the CPI Class I variables showed a marked and consistent tendency in both districts for pupils of higher ability to obtain more favorable scores. It may be concluded that ability grouping in both treatments is positively related to scores on the Class I variables of the CPI with this relationship being somewhat more pronounced for pupils in ability grouped classes. Between-sex comparisons on these variables were generally non-significant although a slight trend favoring girls' groups appeared to be present.

Between-treatment comparisons in the broad area of socialization, maturity and responsibility--the Class II measures on the CPI--showed no consistent pattern of significant differences, excepting for a slight tendency for District A superior pupils to obtain more favorably scores than comparable District R pupils. These results suggest the conclusion that the ability grouping and random grouping treatments have no differential effect upon the broad area of socialization, maturity, and responsibility as measured on the CPI.

Within-treatment comparisons on the Class II variables showed a significant tendency in both districts for pupils of higher ability level to obtain more favorable scores on these six personality characteristics. As was the case with the Class I variables, level differences were more pronounced among ability grouped pupils. The great majority of significant within-treatment differences in both districts were on four of the Class II variables: Responsibility, Socialization, Tolerance and Communality. These results lead to the conclusion that the above variables are significantly related to ability level in both grouping treatments with this relationship being somewhat more pronounced in the ability grouping treatment. A consistent pattern of significant sex differences on the Class II CPI variables favored girls, leading to the conclusion that girls obtain significantly more favorable scores in the broad personality area of socialization, maturity and responsibility than do boys during the junior high school years. These results are generally in line with the findings in adolescent psychology regarding the development of the two sexes during middle adolescence.

Class III variables on the CPI include three scores aimed at measuring Achievement via Conformance (Ac), Achievement via Independence (Ai), and Intellectual Efficiency (Ie). Between-treatment differences on these three variables for subjects in Samples VI and VII were generally non-significant. There was some tendency for superior District A boys to obtain more favorable scores than comparable District R boys on these variables with two of the six comparisons between these groups reaching statistical significance and the remaining four differences following this same trend. With the possible exception of superior boys, it is concluded that the grouping treatments had no influence upon the broad area of achievement potential and intellectual efficiency as measured on the CPI.

Within-treatment comparisons on the Class III variables revealed a marked and consistent tendency for pupils of higher ability in both districts to obtain more favorable scores, with this relationship somewhat more pronounced in District A. It may be concluded that achievement potential and intellectual efficiency are positively related to ability level in both ability grouped and random grouped classrooms.

Of the 36 between-sex comparisons on Class III CPI variables, 20 were statistically significant,all favoring girl's subgroups. Significant sex differences were found in both districts and at all ability levels although significant differences were somewhat more frequent at the superior and average levels. It may be concluded that junior high school girls, especially of the superior and average ability levels, are consistently higher than comparable boys groups on the CPI measures of achievement potential and intellectual efficiency.

Class IV of the CPI is concerned with intellectual and interest modes and is made up of three subtests: Psychological-mindedness (Py), Flexibility (Fx), and Femininity (Fe). No consistent between-treatment differences emerged on any of these three scores leading us to the conclusion that ability grouping and random grouping have no differential effect upon the Class IV CPI variables. There is also a general lack of significant within-treatment differences on the Class IV variables indicating that no consistent relationships between these variables and ability was present in the Utah study. The only noteworthy between-sex differences emerging on the Class IV variables was a fairly consistent tendency for boys to obtain higher scores on Psychological-mindedness than comparable girls' groups.

## REFERENCES

Bettelheim, B. Segregation: new style. Sch. R., 1958, 66, 251-272.

Cattell, R. B. Handbook for the objective-analytic personality test batteries. Institute for Personality and Ability Testing, Urbana, Illinois, 1955.

Drews, E. M. The effectiveness of homogeneous and heterogeneous ability grouping in ninth grade English classes with slow, average, and superior students. Unpublished manuscript. Michigan State University, 1962.

Gough, H. G. Manual for the California Psychological Inventory. Palo Alto, California: Consulting Psychologists Press, 1957.

Gowan, May Seagoe, Why homogeneous grouping? Calif. J. sec. Educ., 1955, 30, 22-28.

Guilford, J. P., & Martin, H. G. The Guilford-Martin Inventory of Factors GAMIN. Manual of Direction and Norms, Beverly Hills, California: Sheridan Supply Company, 1948.

Guilford, J. P., & Guilford, R. B. Personality factors D, R, T, and A. J. abnorm. soc. Psychol., 1939, 34, 21-36.

Hamalainen, A. E. Method of grouping pupils should provide normal social situations. Nations Sch., 1950, 45, 34-35.

Jackson, G. T. Each according to his ability. Sch. Exec., 1943, 62, 37-38.

Kvaraceus, W. C., & Wiles, Marion E. An experiment in grouping for effective learning. Elem. Sch. J., 1938, 3, 264-268.

Luchins, K. B., & Luchins, Edith. Children's attitude toward homogeneous grouping. J. genet. Psychol., 1948, 72, 3-9.

Mann, H. How real are friendships of gifted and typical children in a program of partial segregation? Except. Children, 1957, 23, 199-201.

Marsh, D. E. An interesting experience in grouping. Calif. J. sec. Educ., 1955, 30, 49-59.

Mathias, R. Methodology and results of personality appraisal with Roscharch, TAT and figure drawing. Paper presented at American Psychological Association convention, Cincinnati, Ohio, 1959.

Mosier, C. I. A factor analysis of certain neurotic tendencies. Psychometrika, 1937, 2, 263-287.

Riley, Fay G. Grouping gives each child a chance. Nations Sch., 1956, 58, 51-55.

Simpson, R. E., & Martinson, R. A. Educational programs for gifted pupils. Sacramento, California: California State Department of Education, 1961.

Taylor, C., & Combs, A. W. Self-acceptance and adjustment. J. consult. Psychol. 1952, 16, 89-91.

Tonsor, C. A. Must we segregate? Nat. Assn. Sec. Sch. Prin. Bull., 1953, 37, 75-77.

## Chapter 13

# PERSONALITY-THE PROJECTIVE STUDIES

by

Walter R. Borg, Luna Brite, and Clarence Swaner[1]

## INTRODUCTION

In addition to obtaining data in the Utah study on self-report personality inventories, two aspects of the study were concerned with the measurement of specific personality variables using a projective approach. The first of these studies employed an individually administered measure consisting of eight pictures selected from the Thematic Apperception Test (TAT) (Murray, 1943) and the Michigan Picture Test (MPT) (Michigan State Department of Mental Health, 1953). This study was concerned primarily with comparisons between scores of aggression, depression, and inferiority feelings obtained from slow fifth grade pupils in Districts A and R, and will be referred to as the fifth grade study. This study was carried out by Dr. Luna Brite. The second study adapted substantially the same projective measure for group administration. This study involved sixth grade pupils at all three ability levels and included two measures of anxiety in addition to the projective measure. This project, which will be referred to as the sixth grade study, was carried out by Dr. Clarence Swaner.

## WORKING HYPOTHESES

Research literature available in this area, which was reviewed in Chapter 12, did not supply many clues as to the possible effects of ability grouping on the personality variables to be studied. The reader will recall that Mathias, (1959) found that in heterogeneous classrooms, superior and slow learners showed more signs of aggression and depression than did average pupils. Although his study contained no comparable data on ability grouped children, it had implications for ability grouping as he concluded that superior and slow learning children showed more aggression and depression than the average children because the educational program was not designed for them. This would infer that in an ability grouped situation where the program was more closely adapted to the pupil's ability level, less aggression and depression would be found.

Luchins and Luchins (1948) reported that children in ability grouped classes felt stigmatized by being placed in the low ability group. The inference here is that feelings of inferiority were increased by ability grouping. However, this study lacked a comparable group from heterogeneous or random grouped classrooms so it is possible that similar feelings of inferiority may have developed in low ability pupils in random grouped classes, although perhaps for different reasons.

---

[1] This chapter is based upon the doctoral dissertations of Dr. Brite and Dr. Swaner that were carried out under Dr. Borg's supervision.

The following working hypotheses, which refer to the fifth and/or sixth grade studies as indicated, are based for the most part upon the findings of the above studies. It was hypothesized that:

1. More evidence of aggression would be found in slow learning pupils who were randomly grouped than in slow learning pupils who were placed in ability groups (fifth and sixth grade studies).
2. There would be more evidence of depression in slow learning pupils who were randomly grouped than in slow learning pupils who were ability grouped (fifth and sixth grade studies).
3. There would be more evidence of inferiority feelings in slow learning pupils who were ability grouped than would be found in slow learning pupils who were randomly grouped (fifth and sixth grade studies).
4. Superior pupils in random grouped classrooms would obtain higher (less favorable) mean scores on aggression and depression than similar pupils in ability grouped classrooms (sixth grade study).
5. Ability grouped superior pupils, as a result of intragroup competition and fears associated with losing social status if not successful, would obtain higher manifest and test anxiety mean scores than randomly grouped superior pupils (sixth grade study).
6. Within-treatment comparisons in the sixth grade study would reveal the following differences:
    a. Superior and slow pupils in both grouping systems would obtain higher means scores on the personality variables on anxiety, aggression, and depression than would average pupils.
    b. There would be a positive relationship between the manifest anxiety scores and pupils' ability level, with superior pupils in both grouping situations obtained the higher scores.
    c. Slow pupils would obtain higher anxiety scores than average or superior pupils in both districts.
7. With regard to sex differences, it was hypothesized that:
    a. Boys at all ability levels would obtain higher mean scores on aggression than girls (fifth and sixth grade studies).
    b. Girls at all levels would obtain higher scores on depression than boys (fifth and sixth grade studies).
    c. Girls at all levels would obtain higher manifest anxiety scores and test anxiety scores than boys (sixth grade study).
    d. Girls at all levels would obtain higher mean lie scores than boys on the *Children's Manifest Anxiety L Scale* (sixth grade study).

# PROCEDURE

## The Measures

### Introduction

When this investigation was first considered, it was thought that individually administered TAT-type testing would best serve the purpose at hand, because of its basic rationale:

> . . . the tendency of people to interpret an ambiguous human situation in conformity with their past experiences and present wants, and the tendency of those who write stories (or tell them) to do likewise: draw on the fund of their experiences and express their sentiments and needs, whether conscious or unconscious. (Murray, 1943, p. 1)

The tendency of individuals to interpret present stimuli according to their own accumulated apperceptive mass (based on part experience and present needs) has prompted Bellak (1950, p. 11) to prefer the term "apperceptive testing" rather than "projective testing." Both terms imply that in a comfortable story telling situation, the story teller projects his own inner dynamics into the "social situation" staged by the ambiguous stimuli. The individual's process of interpreting the present stimuli according to experience (appreceptive mass) and present needs is considered to be largely unconscious.

There have been thousands of studies on various aspects of thematic apperceptive testing. There are 610 references on the TAT alone in the Fifth Mental Measurements Yearbook (Buros, 1959).

Rabin, in his symposium article, expresses well one of the reasons for the extensive adaptations of projective testing:

> When research workers are concerned with answers to specific theoretical questions, or with the exploration and elucidation of specific personality variables . . . the traditional techniques have been found wanting. They had to cast about for methods, although based on the projective hypothesis, that are particularly suited to the problems with which they were concerned. (Rabin, 1961, p. 6)

Among the studies that required adaptations of the thematic apperceptive method, are those of Franck and Rosen (1949) who developed their own projective test on masculinity-femininity; Blum (1949), who devised the Blacky Pictures to evaluate psychosexual development; Davids (1955), who modified the usual procedures in his study of alienation; and McClelland, Atkinson, Clark and Lowell (1953), who used a limited number of cards in studying the achievement motive and other variables. Kagan (1956) and Lesser (1958), for their research on aggression, developed their own pictures which were pertinent to the groups being tested (boys 6 to 10, and 10 to 13) and to the personality variable under study. Bellak (1954, 1960) developed the Children's Apperception Test for younger children; Rosenzweig, et. al. developed the Picture Frustration Tests (1948); and the Michigan State Department of Mental Health (1953) developed the Michigan Picture Test for school age children. These studies indicate the growing tendency to reduce ambiguity in favor of selecting stimuli "which tap the specific situation area in which the behavior to be predicted occurs" (Vernier, Whiting, and Meltzer, 1955, p, 181-182).

## The Individual Projective Test

In the Utah study, the investigators wanted the spontaneous reactions of the children to pictures that portrayed situations relatively pertinent to the child and which were selected to reveal the variables to be measured in this research. However, before making the final decision to use TAT-type testing, the literature was searched, faculty members in the Department of Psychology at Utah State University were consulted, and correspondence was carried on with investigators who were involved in adapted TAT-type testing.

There were many tests and studies that were more or less pertinent to the needs of this study, but none of the tests completely met the specific testing needs, nor did any of the scoring methods serve as a model to follow. The final decision of the investigators was to select appropriate pictures from the variations of the TAT and other apperceptive measures available.

Cards from the various measures were evaluated by the investigators and members of the Department of Psychology at Utah State University. After considerable discussion and comparison, nine cards, from the Thematic Apperception Test (TAT) (Murray, 1943) and the Michigan Picture Test (MPT) (Michigan State Department of Mental Health, 1953) were selected for use in the pilot study. The cards chosen are described in Table 13-1. Scoring procedures based on a five point rating scale independently applied by two raters were developed in the pilot study. Procedures in the fifth grade study called for individual administration of the measure with the entire testing session tape recorded. Typed transcripts of the tape recorded protocols were scored independently by the two raters. During the fifth grade study, interrater reliability coefficients were regularly calculated on the blocks of protocols which had been scored between scorer conferences. This served as an objective check of the scorers' adherence to scoring rationale and procedures, and prevented them from drifting apart in their procedures and operational definitions. The composite rater reliabilities for the 144 protocols scored in the fifth grade study were: aggression, .95; depression, .86; and inferiority feelings, .90.[1]

## The Group Projective Test

The sixth grade study, although building to a large extent upon the experience gained in the fifth grade study, required several procedural changes to adapt to the group testing situation.

One new stimulus picture was employed, and the total number was reduced to seven. The pictures were presented to groups of sixth grade pupils by means of an opaque projector in a semi-darkened room. Pupils were required to write seven-minute stories for each of the seven TAT-type pictures, instead of responding orally as in the fifth grade study. Pictures used in the sixth grade phase of the experiment are indicated in Table 13-1.

Having pupils write stories to TAT pictures instead of recording their responses according to the standard procedure considerably shortened the length of the stories

---

[1] Corrected using Spearman-Brown prophecy formula.

Table 13-1

TAT and MPT Cards Used in Utah Study

| Story | Studies Using Picture[a] | Test | Picture Number | Subject of Picture |
|---|---|---|---|---|
| 1 | P56 | TAT | 1 | Boy and violin |
| 2 | P5 | MPT | 6 | Children playing checkers |
| 3 | P56 | TAT | 3 BM | Seated figure leaning on couch |
| 4 | P56 | MPT | 3 | Classroom in session |
| 5 | P | MPT | 2 | Boy and girl in country |
| 6 Boys | P5 | MPT | 10 B | Adult authority figure and child |
| Girls | P5 | MPT | 10 G | |
| 6 Boys | P56 | MPT | 8 B | Boy absorbed in thought |
| Girls | P56 | MPT | 8 G | Girl absorbed in thought |
| 7 | P6 | TAT | 7 GF | A young girl with an older woman |
| 8 | P5 | MPT | 9 | Storm, lightning |
| 9 | P56 | MPT | 12 | Blank card |

[a] P = pilot study; 5 = fifth grade study; 6 = sixth grade study.

elicited, and offered the raters fewer clues upon which to base a score. The investigators felt, however, that if the TAT is an instrument capable of eliciting meaningful samples of a pupil's personality, and if a pupil really projects into TAT stories the sum total of his feelings, than a story of any reasonable length would contain the elements of his personality adjustment.

The scoring method developed for the fifth grade study was applied to the sixth grade sample of 338 TAT-type story protocols, with the same two investigators scoring for both studies. The resulting interrater reliability coefficients obtained for the sixth grade sample were: aggression, .91; depression, .70; and inferiority, .73. These are lower than obtained in the fifth grade study, probably because of the shorter protocols obtained from the group administered test.

## Scoring the Projective Measure

Over the past 25 years, there have been many approaches to the scoring and interpretation of the TAT stories. The original techniques of Murray (1938) and his coworkers were based on the Need-Press method: the needs of the hero (the individual in the story with whom the subject is more closely identifying), and the environmental forces (press) that he is exposed to. Other processes for evaluation of the TAT have varied from very long psychoanalytical analyses requiring 10 or more hours, to very simple counting of words which had been selected as designating the presence of a personality variable being studied.

Since this study was to require the evaluation of nearly 150 protocols in the fifth grade phase, and over 300 in the sixth grade phase, long clinical analysis was inappropriate. Equally inappropriate for evaluating such a rich source as TAT protocols, would be the opposite extreme of a simple word-count technique, for the investigators hoped to tap the inner dynamics as related to the three personality variables being studied. The method of evaluation had to be such that the groups could be compared statistically, however, so some type of quantification was necessary.

There has been much research and disagreement on the problems related to quantifying projective material. Lesser (1961) discussed the pros, cons, techniques, and problems of this controversial and complicated topic in his recent symposium article. The methods of quantification have been quite varied. Some use pertinent word counts; some count, and at time scale, pertinent behavior, usually of the hero; some count stories with pertinent themes; some confine analysis to hero activity; some confine analysis to manifest content; eliminating all inference; some dichotomize while others use intensity scales varying from 3-points to 11-points; some use Q-sort techniques and some use Guttman's scaling.

In this project an attempt was made to retain as much of the holistic approach as possible by treating each story as a whole, and making a clinical, partially subjective evaluation of the story, with the aid of the scoring "helps" that were developed in the pilot study. For quantification, a five-point intensity scale was used which, according to Eron provides:

> . . . an opportunity to combine qualitative, clincial judgements and a more or less rigorous quantification. It is possible for the ratings to be as subjective and holistic as necessary but it is essential that the criteria be verbalized so that the methods can be communicated and subsequent raters can utilize the scales with comparable results. (Eron, 1955, p. 126)

Essentially the same scoring procedures were used in both the fifth grade and the sixth grade studies. In the fifth grade study, the stories were recorded on tape and later typed in duplicate copy. All identifying information except the sex of the pupil was removed from the protocols and the raters had no knowledge of the pupil's district or grade level. In the sixth grade study, the stories were scored directly from the pupils written response, after identification had been removed. Maintaining the anonymity of the pupils was considered necessary to reduce possible scoring bias. The protocols were randomized by shuffling and then scored in units of 15 to 30 per week, according to the established method and utilizing the "helps" that had been developed. During the scoring of these tests, the "helps" lists were continually supplemented as new meaningful material emerged. During the process of scoring, the raters met weekly to discuss methodology, revise the "helps" lists, and discuss any problems that may have been encountered.

In an effort to maintain comparable scoring procedures, the raters frequently would select a previously scored test which had presented scoring problems and discuss the scoring procedure for that protocol in detail. In spite of disagaeements, original scores were never altered and only previously-scored protocols were discussed. Test protocols were always rated independently be each rater and comparisons always followed the recording of the original scores.

## Anxiety Measurement in the Sixth Grade Study

One of the personality variables investigated in the sixth grade study was anxiety. The term "anxiety" has been widely used with a rather large variety of meanings. In what appears to be its most general meaning and most common usage, "anxiety" refers to an affective state of the individual--characterized by fearfulness, apprehension, or dread. Following the original suggestion of Freud, many workers have found it both theoretically and practically useful to distinguish between a generalized state of apprehensiveness, in which the individual's anxiety is not restricted to any particular object or class of objects, and fearfulness or apprehension attached to particular objects or situations. Although the custom is not universal, research workers are tending increasingly to designate the specific types of anxiety that their instruments purport to measure, such as test anxiety, audience anxiety, and sex anxiety.

In addition to determining the specific type of anxiety which an instrument will presumably measure, it is also of value to define the particular behavioral function that the anxiety performs. Mussen (1960) suggests that anxiety must be evaluated both as a motive and also as a response tendency. Response tendency is described as the dispositional quality a subject brings to all situations; and motive, as the pattern or response utilized by the individual that serves the function of reducing, eliminating, or avoiding anxiety. In any appraisal of anxiety in children, one might then ask two basic questions about the anxiety manifested by a child: (a) How anxious and fearful is the child? and (b) To what extent and in what way does his anxiety serve a motivating function in his behavior?

The two tests of anxiety used in this study; The _Children's Manifest Anxiety Scale_ (_CMAS_) and the _Test Anxiety Scale_, (_TASC_) are measures of self-reported anxiety and response tendency (situational anxiety). The tests are objectively scored and have the advantage of being group administered.

## The Children's Manifest Anxiety Scale (CMAS)

The CMAS is a widely used questionnaire measure of children's anxiety. The test was developed by Castaneda, McCandless, and Palermo (1956a) for use primarily with fourth, fifth, and sixth grade children. This 53 item questionnaire represents a childrens' form of the Taylor Manifest Anxiety Scale (TMAS) that was developed by Taylor in 1951. The TMAS in turn was derived from the Minnesota Multiphasic Personality Inventory. The questions deal with reported physiological and psychological concomitants of generalized anxiety such as "I blush easily," "My hands feel sweaty," "I get angry easily," "I often do things I wish I had never done." Taylor (1956) has given us some insight regarding construction of the TMAS. She points out that it was designed to investigate the role of drive (Hullian concept) in certain learning situations--the assumption being that anxiety scores are in some way related to drive level. The construction of the test was not aimed at developing a clinically useful test which would diagnose anxiety, but rather was designed solely to select subjects differing in general drive level.

The CMAS, being derived from the TMAS scale, was also aimed primarily at the measurement of general drive level, or strength of motivation, in experimental studies of the relationship between drive strength and performance in children. In general, the results obtained with the CMAS suggest that it is tapping meaningful anxiety manifestations in children with some validity, and it may therefore be used with profit as a research instrument. Because of the original interest in the scores as measures of drive strength, the criteria against which the test has been validated have not represented a very broad range of theoretical constructs about anxiety as such. Despite the inference that the CMAS anxiety scores are primarily drive level scores, it is interesting to note the overlap between the CMAS, the TASC, and the Audience-Anxiety Scale (AAS)--the latter two being accepted as measures of response (situational) anxiety. Mussen (1960) indicated that intercorrelations among the three tests, based on 421 children in the third to tenth grades, were: AAS, .74 with the TASC, .62, and with the CMAS; TASC with CMAS .70. This suggests considerable overlap in the basic constructs underlying the three tests and suggests that while the CMAS was designed to measure drive level, it is also tapping some of the same areas as the TASC--a fact that might help account for some of the similarities found in the research with each test.

The CMAS Lie (L) Scale. The CMAS contains an 11-item Lie Scale (L) as a measure of response set. The Lie Scale is assumed to reflect the tendency for any given subject to falsify his response to the anxiety text items. The lie score is high for children who deny that rather universal truths are characteristic of them, such as; "I do not lie," "I am always good," and "I never get angry." The initial standardization data presented by Castaneda, McCandless, and Palermo (1956a) indicated that the Lie Scale measures something different from manifest anxiety. Attempts to relate the anxiety scale and the L Scale for different grade levels and for boys and girls separately, produced correlations clustering around zero. As with the anxiety scales, there was a general tendency for girls to receive higher scores in comparison to boys.

Most researchers who have used the CMAS have not used the L Scale, as the items have been demonstrated to have little in common with the direct measurement of manifest anxiety. Eliminating subjects with high lie scores from a study would also eliminate a good part of many experimental groups. McCandless and Castaneda (1956) in one study, and Castandeda, Palermo, and McCandless (1956b) in another study, stated that eliminating subjects with L scores above seven would have reduced the size of their sample N's about 30 per cent.

In the present study, the experimental samples were sufficiently large to allow for the elimination of subjects with L scores over seven and still maintain a sufficient number of cases. The rationale for doing this was that it seemed desirable to work with subjects who were at least willing to attempt telling the truth and who were motivated enough to take an interest.

## The Test Anxiety Scale for Children (TASC)

Mandler and Sarason (1952) developed this 30 item "yes-no" questionnaire which is concerned with attitudes toward and experiences in test and test-like situations. Like Taylor's scale, the TASC is intended to measure anxiety as a personality variable. The hypotheses that have been tested with the TASC focus for the most part upon anxiety as a response tendency, since the test questions deal primarily with affective, physiological, and motor manifestations of anxiety, such as: "Do you worry a lot before you take a test?" "When the teacher says she is going to find out how much you have learned, does your heart begin to beat faster?", and "When you are taking a test, does the hand you write with shake a little?"

The TASC is built around three primary hypotheses, and the research reported in the literature is devoted for the most part to the testing of these hypotheses, which Sarason (1960) and his associates state as follows:

> A. The reaction of the test anxious child to test and test-like situations in the classroom reflects his experience in psychologically or interpersonally similar situations in his home both before and after the beginning of formal schooling.
>
> B. The test anxious reaction has, in addition to its conscious significance, a concurrent unconscious significance which primarily relates to what has been experienced by the child in the family situation. . . . (p. 13)
>
> C. The rate of increase in performance of the test anxious child would be significantly less than for the non-test anxious child. (p. 23)

Sarason and others have administered the TASC to large numbers of school children from grades one to five, and have diligently explored the relationship between their test anxiety measures and a wide variety of psychologically relevant criteria. Mussen (1960) indicated that the overall results of the research done with the TASC questionnaire suggests that it is tapping a meaningful anxiety manifestation in elementary school children.

Sarnoff, Sarason, Lighthall, and Davidson (1958) demonstrated the TASC has satisfactory reliability by administering the test to 100 randomly selected school children in grades one, two, and three. The split-half coefficient of reliability was .81. Sarason and his associates (1958) also cite reliability data that are essentially the same.

The TASC seems well suited for research concerned with anxiety as a function of ability grouping practices. Sarason and his collaborators (1960), as the premise for their first basic hypothesis, point out the relationship between children's feelings of anxiety and parental attitudes and influence. It is apparent in their writings that the child brings to the test and/or school situation the basic predisposition for anxiety, which then may be elicited by the school and/or test environment. It is significant that Sarason states in another reference:

> In our conception, test anxiety is not something that is created in the school, it is something that the child brings, more or less to the school, and in that situation can be either made higher or lower or kept at the same level. (Sarason, 1961, p. 23)

In relation to the present study, this might suggest that children raised in similar cultural and economic environments might, due to the homogeneity of the environmental influences, manifest similar parent-induced anxiety predispositions which might then be brought to the classroom. If these assumptions are valid, it would appear that differences in test anxiety scores between the two grouping treatments used in this study would likely be an effect of the school situation.

## The Sample

### Fifth Grade Study

Subjects for both the fifth and sixth grade studies were selected primarily from Sample IV of the Utah study. In the fifth grade study pupils were selected from four schools in District A. All pupils in the ability grouped slow sections of these schools, a total of 41 boys and 32 girls, were administered the individual projective measure. In District R, pupils were selected from six schools. A sufficient number of slow pupils were not available in the Sample IV and as a result additional District R cases were added based on the opinions of teachers and previous achievement records. The mean achievement level of District R slow pupils at the end of the fifth grade study was significantly higher than the level of the District A pupils.[1] In District R a total of 73 pupils, 39 boys and 34 girls were tested. Of those pupils who were administered the individual projective measure in the two districts, two protocols were lost due to faulty tape recording leaving a final overall sample of 144 pupils in the fifth grade study.

### Sixth Grade Study

The experimental design of the sixth grade study called for a sample of 180 pupils from each of the two grouping treatments, for a total of 360. Sixty pupils (30 boys and 30 girls) at each ability level were to be studied but a shortage of slow girls in both districts, made it impossible to obtain sufficient cases in these subgroups (see Table 13-2).

As the District A pupils were already grouped within their respective schools, it was a simple matter to select enough schools and test enough classes to obtain the desired number of cases in each group. In general, an effort was made to test all of the sixth grade pupils in each of the selected schools. However, as two of the larger schools had several average groups, only one average section was tested

---

[1] These pupils were comparable at the beginning of grade four. Differences reflected greater gains for District R pupils.

Table 13-2

Sample Composition in the Sixth Grade Study

| | Ability Level | | | | | | |
|---|---|---|---|---|---|---|---|
| | Superior | | Average | | Slow | | Total |
| | Boys | Girls | Boys | Girls | Boys | Girls | |
| District | | | | | | | |
| A | 30 | 30 | 30 | 30 | 30 | 13 | 163 |
| B | 30 | 30 | 30 | 30 | 30 | 25 | 175 |
| | | | | | | | 338 |

in each of these schools. All superior and slow groups were tested in each of the selected District A schools.

Obtaining the desired sample in the District R schools presented other problems. As the pupils were grouped randomly, it was sometimes necessary to test an entire class in order to obtain data on one or two slow or superior children. As a result, it was necessary to test 707 District R pupils in order to obtain a sample with enough cases of each sex at each ability level. After all of the data had been collected, the protocols were separated into 12 groups on the basis of treatment, sex, and ability level. Each protocol was checked for completeness and those protocols in which a pupil failed to write a story on a particular picture stimulus or wrote an incomprehensible story were rejected. The CMAS Lie Score was also utilized as a method of screening protocols and all pupils who obtained L scores over seven were eliminated from the sample. During the course of testing the pupils were observed carefully and the test protocols of those few pupils who obviously made no effort on the test were eliminated from the sample. Several pupils of low ability could neither read nor write and protocols of these pupils were not used although the pupils were allowed to remain in the test situation in an effort to avoid disturbing the class environment. The application of the above screening procedures resulted in the rejection of a considerable number of test protocols. Following screening, 30 protocols were randomly selected from each of the 12 groups with the exception of the slow girls in both districts. All scorable protocols were used for these groups because of the limited number available.

## The Pilot Studies

### Fifth Grade

Prior to the start of the fifth grade study, a pilot study was planned and carried out. Among the purposes of the pilot study were the development of techniques for administering the individual tape recorded projective measure to fifth grade children in a school setting; development of forms for recording test protocols, scores and interview data; the establishment of a scoring rationale, and a preliminary evaluation of the effectiveness of the nine TAT-type cards that had been selected.

The pilot study sample included 26 pupils in the fifth grade class at the Edith Bowen Laboratory School of Utah State University. This class was considerably more heterogeneous than the slow pupils to be used in the main study, but because of this provided a wider range of responses to the projective material than could have been obtained from a more restricted group. The projective measure was administered individually to pilot study subjects in the faculty room of the laboratory school. All testing and interviewing was tape recorded, and pilot study protocols were typed from the tape recordings. The design called for two persons to independently score each test to be included in the main study in order to make it possible to estimate inter-rater reliability. The two test scorers worked together on the first few pilot study protocols and developed a five-point rating scale for each of the three personality variables to be measured. Simultaneously, a list of words, ideas, and story situations was developed for each personality variable which contained items that indicated the presence of the variable in fifth grade children. These, lists, designed to help the scorers evaluate subsequent protocols were called "helps" (see Appendix G). These "helps" were expanded, revised, and refined throughout the study. During the pilot study, a set of procedures for scoring the tape recorded protocols was also established.

In general, manifest content was scored; but in cases where doubt existed, the scorer was free to use his subjective impression based on inferences drawn from the total test protocol.

The procedure established during the pilot study for evaluation of each story was as follows:

1. A typed transcript with pupil identification removed was presented to each rater in the same order.
2. Each story was read twice by each rater.
    a. The first reading involved getting the total gist of the story and assigning tentative numerical ratings for each variable.
    b. The second reading required synthesizing the diagnostic impression and rating each variable according to the five-point scale.
3. Scores were recorded on a score sheet which permitted tabulating the scores for each story as well as a total score for each personality variable.

Once the various scoring tools had been established, the remainder of the pilot study protocols were scored independently by the two scorers in order to obtain an estimate an inter-rater reliability. Reliabilities on the 19 pilot study protocols scored independently, corrected using the Spearman-Brown prophesy formula were: aggression, .90; depression, .83; and inferiority, .78.

Sixth Grade

In preparation for the sixth grade study, it was decided to conduct a brief pilot study on a group of pupils similar to those to be used in this phase of the project. The purposes of this trial were to try out the group projective and anxiety measures and determine the best order of presentation, test the utility of projecting TAT pictures by means of an opaque projector in the usual classroom situation, determine how many TAT pictures would be needed to yield satisfactory data, determine how much time pupils would need to write their stories, score a sample of written stories in order to check the adequacy of the scoring method developed for the fifth grade study, determine if sixth grade pupils could write scorable stories, and try out and refine administrative procedures.

One of the schools in District A not involved in the Utah study was selected for the pilot study and the superior, average and developmental sections were given the test battery. It had been decided that, whenever possible, pupils should be tested in their own classroom situation and that different classes would never be combined for the purpose of administering the test. This was done in an effort to reduce the possibility of introducing intervening variables such as anxiety or hostility which might arise in a strange situation.

The results of the pilot study indicated that in general sixth grade pupils could write adequate stories to as many as seven TAT pictures, but became restless and disinterested when more than seven pictures were used. Five minutes writing time per card was found to be too short for the superior pupils, and many of the average pupils also expressed a need for a "little more time." Most of the slow pupils were finished in five minutes. Ten minutes writing time either fatigued most pupils or left time for "monkey business" between card administrations. Seven minutes writing time was tried and proved satisfactory in all situations, although occasionally a pupil did not have an opportunity to complete his story.

As the TASC instructions contain an introduction to the test situation, it was given first. This was followed by the projective measure (seven cards), and the CMAS was given last. It was found that the test battery could be given in 90 minutes, allowing for a three-minute break after the fourth TAT card.

All of the pilot study sessions were tape recorded, and an effort was made to structure the final test instructions so that most of the questions commonly asked by the pupils were answered in the instructions. After each trial, the class was encouraged to offer suggestions that would make the test better and to describe their feelings about this kind of test. In general, most of the children considered the test a new experience and many stated that they "wished more tests were like this."

## RESULTS

### Fifth Grade Study

The pupil's composite score was obtained for each of the three personality variables by combining the scores given him by the two independent raters. In order to determine whether significant differences between the grouping treatments were present, the mean scores of ability grouped and random grouped boys and girls were compared using the t-test. Table 13-3 summarizes the difference between District A and District R pupils on the three personality variables for which the test protocols were scored. It may be seen that none of these differences reached statistical significance. In comparing the District A and District R pupils used in this phase of the study on the Sequential Tests of Educational Progress (STEP) it was found that the District R pupils had a higher composite STEP mean than District A pupils. As the STEP scores were found to correlate to some extent with scores on the three personality variables, there was a possibility that this difference in STEP means could have influenced the means of the personality variables sufficiently to have altered the results obtained using t-tests. Thus, the between-treatment personality data were reanalyzed to adjust for the initial STEP differences using analysis of covariance. None of the between-treatment differences were statistically significant on this analysis.

Sex differences on the three personality variables are summarized in Table 13-4. It will be noted that boys made somewhat higher aggression scores in both ability grouped and random grouped classroom with this difference reaching statistical significance for District A pupils and approaching significance for District R pupils. On the depression scores, girls were higher in both districts with this difference reaching significance at the .01 level for District R girls. Sex differences as measured by the projective test on inferiority feelings were not significant in either district.

Four subtests from the Sequential Tests of Educational Progress in the areas of mathematics, science, reading, and social studies were administered throughout the Utah study to obtain comparative achievement data for ability grouped and random grouped pupils. A composite score on these four STEP measures was available for all District A slow pupils in this study, and for 59 of the 73 District R pupils on whom the projective personality data were collected. Correlation coefficients were computed between the STEP composite score and each of the three personality variables for the four subsamples used in the fifth grade study. These correlations are summarized in Table 13-5. The correlations between the STEP score and aggression was statistically significant for District A boys. This correlation suggests that, within the range of ability present in the District A slow groups, the more able students

Table 13-3

Between-Treatment Differences on the Three Personality Variables Measured in the Fifth Grade Study

| | | | Aggression | | | Depression | | | Inferiority | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | $N_A$ | $N_R$ | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t | $M_A$ | $M_R$ | t |
| Slow Boys<br>Dist. A vs. R | 40 | 39 | 29.6 | 27.8 | NS | 18.6 | 18.1 | NS | 25.8 | 24.3 | NS |
| Slow Girls<br>Dist. A vs. R | 31 | 34 | 25.0 | 25.3 | NS | 19.3 | 20.5 | NS | 26.0 | 26.4 | NS |

Table 13-4

Differences Between the Sexes in the Fifth Grade Study

| | | Variables | | | | | |
|---|---|---|---|---|---|---|---|
| | | Aggression | | Depression | | Inferiority | |
| Groups | N | Mean | t | Mean | t | Mean | t |
| Ability Grouped | | | | | | | |
| Boys | 40 | 29.6 | 2.39* | 18.6 | 0.97 | 25.8 | 0.12 |
| Girls | 31 | 25.0 | | 19.5 | | 26.0 | |
| Random Grouped | | | | | | | |
| Boys | 39 | 27.8 | 1.66 | 18.1 | 3.48** | 24.3 | 1.35 |
| Girls | 34 | 25.3 | | 20.5 | | 26.4 | |

* Significant at .05 level.
** Significant at .01 level.

expressed more aggression than the less able boys. This result may suggest that membership in the low ability class is a frustrating experience for those boys who are near the top of this group.

There was a rather marked difference between the districts in the correlations obtained between the STEP score and depression. Correlations in the ability grouped situation were positive and not significant, while those in the random grouped situation were negative and statistically significant. These results indicate that the lower the ability of the slow pupil in the random grouped situation, the greater was his tendency toward depression. It is interesting to note that the relationship between the STEP score and depression was higher for random grouped boys than for random grouped girls even though the girls had significantly higher mean depression scores than the boys. Although these relationships are low, they could indicate that there is a tendency for pupils at the bottom of the ability distribution to experience more rejection in random grouped classes and consequently suffer more depression.

When the significant positive relationship between scores of STEP and aggression is coupled with the similar significant correlation between STEP scores and inferiority feelings for District A slow boys, it suggests that the ability grouping situation may create an unfavorable climate for the personality development of boys in slow groups who are near the top of the ability range for these groups. The positive correlations between the STEP scores and inferiority feelings suggest that as the pupil's ability rises, there is also an increased insight into his ability and his limitations which may bring about the feelings of inferiority.

In reviewing these correlational data, the reader should remember that although statistically significant, the correlations discussed were all quite low.

## Sixth Grade Study

A total of 36 comparisons were made between similar pupils in District A and District R. Of these, six differences were significant beyond the .05 level. These differences are summarized in Table 13-6.

### Projective Data

It was hypothesized that the superior and slow pupils in the random grouped situation would obtain higher mean scores on the variables of aggression and depression than similar pupils in ability grouped classrooms, as an effect of frustrations imposed by the restrictions of a school situation often considered to be geared primarily to the needs of the average pupil. This hypothesis was based on the assumption, derived from Mathias' (1959) study, that the wide spread of individual differences within the heterogeneously grouped classroom would frustrate both the bright and slow pupils and cause a greater amount of aggression and depression in both groups. The only findings in keeping with this hypothesis was that District R slow girls demonstrated significantly more depression than slow girls in District A. However, the relationship applies only to girls and extends upwards to include the average group. Mathias' inference was that the average group would demonstrate less depression than either of the extreme groups. It is possible that girls of average or low ability, grouped on the basis of ability, feel more comfortable, accepted, and have more opportunities for success than similar pupils in the random grouped situations where direct competition with superior pupils may limit their opportunities to demonstrate adequacy. However, the work of Luchins and Luchins (1948) suggested that

## Table 13-5

## Correlations Between STEP and Personality Variables in Fifth Grade Study

| Group | N | Pearson r | Corrected$^a$ r |
|---|---|---|---|
| **STEP and Aggression** | | | |
| **District A** | | | |
| Boys | 40 | .251 | .347* |
| Girls | 31 | .105 | .160 |
| **District R** | | | |
| Boys | 31 | -.067 | -.106 |
| Girls | 28 | .203 | .330 |
| **STEP and Depression** | | | |
| **District A** | | | |
| Boys | 40 | .080 | .114 |
| Girls | 31 | .198 | .296 |
| **District R** | | | |
| Boys | 31 | -.303 | -.452* |
| Girls | 28 | -.234 | -.375* |
| **STEP and Inferiority Feelings** | | | |
| **District A** | | | |
| Boys | 40 | .390* | .517** |
| Girls | 31 | .218 | .321 |
| **District R** | | | |
| Boys | 31 | .229 | .351 |
| Girls | 28 | .178 | .289 |

*Significant at the .05 level.
**Significant at the .01 level.
$^a$ Corrected for restriction in range.

pupils in lower ability groups should feel more depressed because of feelings of discrimination in an ability grouped situation. On aggression, the only significant difference obtained was in the direction opposite that hypothesized, District A superior boys having scored significantly higher on this variable than comparable District R boys. The fact that the relationship is inverse, and does not hold for the slow pupils, casts serious doubt upon Mathias' interpretation of his findings.

Mathias' and Luchins and Luchins' findings applied to both girls and boys as a group. However, in the present findings a sex differential is present that suggests that any inferences concerning the effects of a grouping practice upon any particular group of school children may have to consider the sex of the pupils as well as their level of academic achievement.

Regarding inferiority feelings, it was hypothesized that ability grouped slow pupils would react to being classified as "dumb" by obtaining higher mean scores on this variable than similar heterogeneously grouped pupils. The findings of the sixth grade study revealed no significant sex differences, boys and girls at each level were combined so as to increase the size of the sample in each group, and the combined groupings were retested for significance. Analysis of combined groupings failed to yield any significant between-treatment differences. Actually, as may be seen in Table 13-6, both boys and girls received somewhat higher inferiority scores in District R, although these differences were well short of statistical significance. Thus, the claim that ability grouped slow pupils would react to the stigma of "being dumb," as advanced by Luchins and Luchins (1948), with increased feelings of inferiority, was refuted. Data from the Utah study suggest that the method of grouping probably is not a significant factor in the development of feelings of inferiority among elementary school children. It might be that the variable of inferiority is a much more stable concept and not as easily influenced as other more "superficial" personality variables.

On the other hand, low ability pupils, grouped as such, may experience an increase in ego strength and feelings of adequacy due to less peer competition and greater chances for success. Ability grouping may also orient the pupil's parents to a better understanding his potential and reduce parental expectation. It is possible that the aforementioned factors may offset any undesirable effect that the "dumb" label may have upon the slow pupil.

## Anxiety Data

There were no significant between-treatment differences in anxiety scores on the TASC or CMAS tests. It had been hypothesized that ability grouped superior pupils would react to the greater intensity of intra-group competition with an increase in anxiety, centered around the fear of losing social status if not successful. The results of the Luchins and Luchins study (1948) suggested that homogeneous grouping practices created a sort of caste system in the school, which tended to make superior pupils overly concerned with their status. Superior pupils in this study are reported to have expressed the feeling that they would rather take a chance of failing in the bright group than be in any way associated with the average or slow groups. The fear associated with losing status might be expected to produce considerable anxiety. Engle (1938) also suggested that bright students may develop increased anxiety as a result of increased expectations and a feeling of being different from their peers. The findings of the Utah study suggest that grouping is not a factor in creating significant anxiety. Significant level and sex differences were found, however, and

Table 13-6

Between-Treatment Differences on the Sixth Grade Study Variables

| | | | Aggression | | | Depression | | | Inferiority | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | $N_A$ | $N_R$ | $M_A$ | $M_R$ | Sig. Level | $M_A$ | $M_R$ | Sig. Level | $M_A$ | $M_R$ | Sig. Level |
| Superior Boys Dist. A vs. R | 30 | 30 | 33.60 | 26.57 | .01 | 17.33 | 17.80 | NS | 31.17 | 29.50 | NS |
| Superior Girls Dist. A vs. R | 30 | 30 | 28.80 | 26.00 | NS[a] | 19.97 | 19.50 | NS | 34.00 | 30.97 | NS |
| Average Boys Dist. A vs. R | 30 | 30 | 28.82 | 27.97 | NS | 18.16 | 18.10 | NS | 31.77 | 30.73 | NS |
| Average Girls Dist. A vs. R | 30 | 30 | 28.43 | 28.23 | NS | 19.23 | 21.97 | .05 | 31.17 | 32.23 | NS |
| Slow Boys Dist. A vs. R | 30 | 30 | 27.97 | 27.43 | NS | 17.43 | 17.83 | NS | 27.67 | 29.23 | NS |
| Slow Girls Dist. A vs. R | 13 | 25 | 26.85 | 27.68 | NS | 16.46 | 19.68 | .05 | 27.46 | 30.80 | NS |

| TASC | | | CMAS | | | CMAS (L)[b] | | |
|---|---|---|---|---|---|---|---|---|
| $M_A$ | $M_R$ | Sig. Level | $M_A$ | $M_R$ | Sig. Level | $M_A$ | $M_R$ | Sig. Level |
| 8.47 | 6.37 | NS | 11.60 | 9.20 | NS | 2.50 | 4.30 | .01 |
| 12.57 | 10.40 | NS | 15.83 | 14.83 | NS | 3.73 | 2.87 | .01 |
| 8.33 | 10.27 | NS | 11.20 | 11.40 | NS | 3.70 | 3.60 | NS |
| 14.57 | 12.90 | NS | 16.13 | 18.40 | NS | 3.33 | 3.23 | .01 |
| 11.40 | 12.67 | NS | 13.27 | 14.07 | NS | 4.17 | 3.80 | NS |
| 14.08 | 11.90 | NS | 17.54 | 15.12 | NS | 5.00 | 4.84 | NS |

[a] Not statistically significant.

[b] CMAS lie scale data were based on all pupils tested, before screening out those with L scores over seven.

these suggest that intergroup and intragroup relationships within the particular grouping system are more important than the effects of grouping or not grouping. These differences will be discussed later in this chapter.

No significant differences were hypothesized on the CMAS Lie Scale. However, three significant between-treatment differences emerged on this variable (see Table 13-6). These findings suggest a between-treatment sex differential between superior pupils in both grouping systems. In District R, superior boys manifested higher L scores than homogeneously grouped superior boys. The finding was opposite for girls, with ability grouped superior girls manifesting higher L scores than random grouped superior girls, thus inferring that the two grouping systems had differential effects upon superior pupils of the two sexes. This in turn suggests that neither method of grouping is necessarily more appropriate for all pupils. The higher L score of superior boys in the random grouping situation may demonstrate more defensiveness. These findings may have relevance when we recall that ability grouped superior boys obtained higher aggression scores than the superior boys in District R. The fact that District R boys are more defensive might help account for the difference in aggression as it is possible that they are suppressing their aggression.

The results suggest that the CMAS Lie Scale may have more discriminative ability and be a stronger tool in the study of personality structure than has been recognized. It is noteworthy that of the six significant between-treatment differences found in the sixth grade study three were on the CMAS Lie Scale. Further research with this variable would probably be of value.

## Ability Level Comparisons

It was hypothesized that superior and slow pupils in both districts would demonstrate higher mean scores on the personality variables of anxiety, aggression, and depression than average pupils. The significant findings relative to this hypothesis are summarized in Table 13-7.

Whenever one works with extreme groups, it might be expected that responses to a stress situation will be more extreme than the responses of an average person in the same situation. However, the findings indicated that pupils of different ability levels in each district reacted somewhat differently. It appears that in regard to the slow pupils, heterogeneous grouping practice tends to increase the anxiety scores of both sexes. Homogeneous grouping practice appears to have its major effect upon superior pupils, leading to higher aggression scores for boys and higher depression scores for girls.

Only one significant ability level difference emerged from comparisons on the CMAS. Slow boys obtained significantly higher CMAS scores than superior boys in the same district, which is contradictory to the "drive level" hypothesis put forth by Taylor (1956). Further, the assumption that there would be a hierarchy of CMAS scores proceeding from superior, through average, to the slow level in descending order was not demonstrated. A review of the mean scored in Table 13-6 shows no consistent relationship between ability level and L scores in either district.

This hypothesis was based on the assumption put forth by Taylor (1956) which suggests that superior pupils as a measure of "drive level," should obtain higher CMAS scores than average or slow pupils. The fact that slow pupils in District R obtained higher scores than superior pupils raises some serious doubts as to the ability of the CMAS to measure drive level. It may be, however, that slow boys

## Table 13-7

## Significant Ability Level Differences

| Ability Level Comparisons | Aggression | Depression | Inferiority | TASC | CMAS | CMAS (L) |
|---|---|---|---|---|---|---|
| **District A (ability grouped)** | | | | | | |
| Boys: | | | | | | |
| Superior vs. Slow[a] | .01[b] | | | | | .01[c] |
| Superior vs. Average | .05 | | | | | .05[c] |
| Girls: | | | | | | |
| Superior vs. Slow | | .05 | | | | |
| Superior vs. Average | | | | | | .01 |
| Slow vs. Average | | | | | | .01 |
| Combined boys and girls: | | | | | | |
| Superior vs. Slow | | | .01 | | | |
| Average vs. Slow | | | .01 | | | |
| **District R (random grouped)** | | | | | | |
| Boys: | | | | | | |
| Average vs. Superior | | | | .05 | | |
| Slow vs. Superior | | | | .01 | .05 | |
| Girls: | | | | | | |
| Average vs. Superior | | .05 | | | | .01 |
| Slow vs. Superior | | | | .05 | | |

[a] The first group listed obtained the highest score; higher scores are considered less desirable.

[b] Level of significance.

[c] Reversed relationship; latter group listed obtained higher score.

exhibit more drive level in an attempt to meet the competition in the random grouping system, while superior boys can keep up with little or no drive. That the CMAS is more a measure of situational anxiety than drive level is suggested by the TASC findings. The TASC, which is purported to be a measure of situational anxiety, did not differentiate between districts, but the ability level differences were essentially the same as those obtained by the CMAS.

The finding that District R pupils of lower scholastic standing demonstrated higher anxiety scores than superior pupils, suggests a greater situational stress upon heterogeneously grouped pupils and, in keeping with research findings, would suggest that lower ability pupils would experience a greater decrement in learning ability than superior pupils due to the effects of increased anxiety. As District A demonstrated no significant district or level differences in CMAS or TASC scores, one might assume that the practice of grouping pupils into ability levels may have the effect of reducing situational anxiety and of contributing to the learning efficiency of the pupil. This would be a valuable area of further research.

It was also hypothesized that slow pupils in both districts would obtain higher test anxiety (TASC) mean scores than would average or superior pupils. The basis for this assumption was that slow pupils in either grouping system would be threatened by any test situation because of poor past performance and would thus demonstrate higher TASC scores. Comparisons between pupils at different ability levels in each district suggested that slow pupils in homogeneously grouped classes are less threatened by test situations than comparable heterogeneously grouped students. The reasons for the difference may be many, but it seems reasonable that the test constructed by teachers for ability grouped pupils might be geared more to their level of ability and offer these pupils more opportunities for success.

Differences between pupils at different ability levels on feelings of inferiority were not hypothesized. Some significant differences on this variable occurred, however, as may be seen in Table 13-7.

These findings suggest that superior and average pupils demonstrate more feelings of inferiority in the ability grouping situation than pupils of low ability in the same grouping treatment. It will be recalled that District A superior boys earned higher mean aggression scores than District A slow boys, and a similar relationship held for District A superior and slow girls on depression. It is generally accepted that in our society boys demonstrate more aggression than girls, and that girls obtain higher depression scores than boys. It might be that superior girls in the ability grouping situation react to feelings of inferiority with depression, while ability grouped superior boys react to feelings of inferiority by demonstrating aggression. These data suggest that homogeneous grouping practices may be producing differences between pupils at the different ability levels that reflect a negative effect upon the superior pupil. It may be that the superior pupils in District A have been adversely affected by the change from an easy competitive situation in their first years of schooling in random grouped classes to the more competitive situation after ability grouping started. This could have a threatening effect upon the superior pupil as he now has to compete with more intelligent pupils. If this aggression is a compensation for feelings of inferiority, and assuming that feelings of inferiority are detrimental to the progress and development of the pupil, the modern-day emphasis upon grouping superior students in order to offer them the "benefits" of an accelerated training may be ill advised. In terms of personality development, grouping, by ability, is of more value to the low ability pupil and of less value to the superior pupil.

## Sex Differences

The only significant sex difference in aggression was that ability grouped superior boys demonstrated more aggression than ability grouped superior girls (see Table 13-8). This finding was in keeping with research findings, except that it should have extended across all levels. The fact that it did not, suggests something inherent in the homogeneous grouping system. Feelings of inferiority have been suggested as the reason for this difference. It was hypothesized that girls at all levels would manifest higher mean scores on the personality variable of depression than boys. The findings on this variable were in the predicted direction, but were significant only for selected groups. In District R, average girls obtained significantly higher depression scores than average boys in the same district, but there were no other significant sex differences. There is no apparent reason why average girls in District R, as a selected group, should obtain higher scores on the variable of depression than average boys in the same district. It is noteworthy that District R average girls also demonstrated higher Manifest Anxiety scores than average boys in the same district, and it is possible that there is something inherent in the grouping system that contributes to this condition. This finding is in keeping with the general observation that girls customarily demonstrate more depression than boys. However, a more complex condition is suggested by the fact that District A superior girls also obtained higher Test Anxiety, Manifest Anxiety, and CMAS (L) scores than superior boys in the same district. Also, one would question why only the superior girls in District A demonstrated the predicted sex difference.

It was hypothesized that girls at all levels would obtain higher mean Manifest Anxiety test scores than boys. This hypothesis held for superior and average pupils in both districts--but did not hold for the slow pupils in either district. However, a comparison of the non-significant mean scores earned by slow pupils demonstrated a trend toward sex differences that were in keeping with the hypothesis. Research has suggested that girls in our culture usually demonstrate higher Manifest Anxiety scores than boys, primarily because our society condones this type of expression for girls, but inhibits such expression in boys by considering such behavioral manifestations as unmanly. The present findings indicate that in this sample of sixth grade pupils, slow boys in both districts demonstrated substantially as much Manifest Anxiety as girls. A possible reason for this might be that low ability boys may have fewer means of demonstrating adequacy simply because of limited ability, and any situation which is likely to evoke a failure could precipitate greater anxiety.

It was hypothesized that girls at all levels would obtain higher mean Test Anxiety scores than boys. Three significant differences on the Test Anxiety scores emerged (see Table 13-8). It will be noted that the significant differences found were all in the predicted direction. Girls at all levels obtained higher mean Test Anxiety scores than boys, but the difference was not significant at all levels (see Table 13-6). These findings demonstrate that while not significant, slow girls in both districts earned consistently higher anxiety scores than slow boys.

Research evidence has indicated that girls customarily obtain higher CMAS (L) scores than boys. In the present study, District A superior and slow girls obtained significantly higher (L) scores than boys at the same levels. The average group demonstrated no sex differences in either treatment. In District R, slow girls obtained higher (L) scores than slow boys but no other significant differences were indicated. This suggests that superior and slow girls in a homogeneous grouping situation tend to falsify their responses to the CMAS more than boys in the same

Table 13-8

Significant Sex Differences on Personality Variables in the Sixth Grade Study

| Sex Comparisons | Personality Variables<br>Aggression | Depression | Combined Inferiority | TASC | CMAS | Total CMAS (L) |
|---|---|---|---|---|---|---|
| **District A** | | | | | | |
| Superior level | .05[b] | .05[c] | | .05[c] | .05[c] | .01[c] |
| Boys vs. Girls[a] | | | | | | |
| Average level | | | | | | |
| Girls vs. Boys | | | | .01 | .05 | |
| Developmental level | | | | | | |
| Girls vs. Boys | | | | | | .05 |
| **District R** | | | | | | |
| Superior level | | | | | | |
| Girls vs. Boys | | | | .05 | .01 | |
| Average level | | | | | | |
| Girls vs. Boys | | .01 | | | .01 | |
| Developmental level | | | | | | |
| Girls vs. Boys | | | | | | .05 |

[a] The first group listed obtained the higher score; higher scores are considered less desirable.

[b] Level of significance.

[c] The relationship is reversed; the latter group listed obtained higher scores than the first.

grouping situation. The tendency towards falsification suggests a response set towards defensiveness and would indicate that the girls at these levels are more concerned with their performance than boys. The fact that few significant differences were obtained in the sixth grade study suggests that the general effects of the two grouping treatments on the variables studied are in most cases similar with the majority of pupils relating to either system in a fairly uniform manner. The absence of clear-cut district differences suggests that except in special cases the specific effect of a particular grouping practice is not significant in terms of the variables studied.

## SUMMARY AND CONCLUSIONS

This chapter deals with two phases of the Utah study concerned with the measurement of personality and anxiety. The first phase, referred to as the fifth grade study, measured aggression, depression, and inferiority feelings on an individually administered TAT type projective measure using a sample of 144 slow pupils taken from ability grouped and random grouped fifth grade classes. The second phase, referred to as the sixth grade study, employed a similar group-administered projective measure as well as measures of anxiety, using a sample of 338 slow, average, and superior pupils taken from ability grouped and random grouped sixth grade classrooms.

Between-treatment comparisons on aggression, depression and inferiority feelings in the fifth grade study showed no statistically significant differences. Between-treatment differences on these three personality variables obtained in the sixth grade study from a group projective measure also failed to show significant differences between slow pupils on these three personality variables with the exception of slow girls in District R who obtained significantly higher scores on depression. District R slow girls in the fifth grade study also showed higher depression scores although the difference was not significant. From the data presented in this chapter, it may be concluded that:

1. Ability grouping and random grouping do not differentially affect the scores of slow pupils on aggression, depression and inferiority feelings excepting for a tendency for slow girls in random grouped classes to show higher levels of depression than slow girls in ability grouped classes. The lack of significant differences in inferiority feelings is particularly noteworthy as these findings support the data obtained on other samples using the inferiority subscore of the GAMIN. Taken together, these results present strong evidence that ability grouping as employed in the Utah study does not cause the development of inferiority feelings among slow pupils.

2. The projective data obtained in the sixth grade study at all ability levels failed to support Mathias' (1959) contention that aggression and depression are higher in random grouped classes for superior and slow pupils because of the frustrations encountered by these pupils in a program geared for the average. Perhaps Mathias' findings would have been supported in random grouped classes where no enrichment provisions were made, but in the Utah study only one significant difference, between slow girls in the two districts supported this contention regarding depression. A significant difference between superior boys in aggression was in the direction opposite the Mathias' hypothesis. As the few significant between-treatment differences found on aggression and depression were not supported by data obtained in the fifth grade study, the most logical conclusion would be that these

variables are not influenced by the ability grouping or random grouping treatments employed in the Utah study.

3. Sex comparisons on the fifth grade study showed boys to be significantly higher in aggression in District A, with this difference approaching significance in District R. Girls were higher on depression in both districts with this difference being significant for District R. These differences, although following the same trend for pupils in the sixth grade study, were not statistically significant. The less adequate TAT stories obtained from slow pupils in the group testing situation may have been the reason for the lack of significant differences in the sixth grade study. In the sixth grade study, superior boys obtained significantly higher aggression scores while superior girls obtained significantly higher depression scores.

Although there were no significant relationships between aggression and depression as measured in this study, both aggression and depression were positively correlated with inferiority feelings. These data suggest that both aggression and depression may at least be partially expressions of feelings of inferiority with boys more frequently expressing their inferiority feelings through aggression while girls more frequently express theirs through depression.

4. No significant differences were found on the TASC or the CMAS, the anxiety measures employed in the sixth grade study. From these results we may conclude that the two grouping treatments have no differential effect upon pupil anxiety at any ability level in either the ability grouped or random grouped classrooms.

5. Correlations computed in the fifth grade study between achievement as measured by the STEP and aggression suggest that the ability grouping situation provides some frustrations for pupils whose ability places them near the top of the slow group. On the other hand, in random grouped classrooms, the correlations between depression and STEP scores suggest that pupils at the bottom of the ability distribution in such classrooms tend to be more depressed, possibly because of their lack of ability to compete in the heterogeneous classroom environment.

6. Level differences in the sixth grade study showed District R slow pupils to obtain higher anxiety scores than superior pupils. These data suggest the tentative conclusion that heterogeneous grouping places a greater situational stress upon pupils of lower ability.

7. Level differences also showed higher aggression scores for superior boys and higher inferiority scores for boys and girls in District A, while no such differences emerged in District R. It may be tentatively concluded that superior pupils in District A were adversely affected by the change from an easy competitive situation in their first years of schooling in random grouped classes to the more difficult competitive situation in ability grouped classes.

## REFERENCES

Bellak, L. On the problems of the concept of projection. In L. A. Abt and Leopold Bellak (Eds). *Projective psychology*. New York: Grove Press, 1950, Pp. 7-32.

Bellak, L. *The Thematic Apperception Test and The Children's Apperception Test in clinical use*. New York: Grune and Stratton, 1954.

Bellak, L., & Adelman, Crusa. The Children's Apperception Test (CAT). In A. T. Rabin and Mary R. Haworth (Eds). Projective techniques with children. New York: Grune and Stratton, 1960, Pp. 62-94.

Blum, G. S. A study of the psychoanalytic theory of psychosexual development. Genet. Psychol. Monogr., 1949, 39, 3-99.

Buros, O. K. The fifth mental measurements yearbook. Highland Park, N. J.: Gryphon Press, 1959.

Castaneda, A. , McCandless, B. R., & Palermo, D. S. The children's form of the Manifest Anxiety Scale. Child Develpm., 1956a, 27, 317-26.

Castaneda, A., McCandless, B. R., & Palermo, D. S. Complex learning and performance as a function of anxiety in children and task difficulty. Child Develpm., 1956b, 27, 327-32.

Davids, A. Generality and consistency of relations between the alienation syndrome and cognitive processes. J. abnorm. soc. Psychol., 1955, 51, 61-67.

Engle, T. L. A study of the effects of school acceleration upon the personality and social adjustments of high school and university students. J. educ. Psychol., 1938, 12, 523-39.

Eron, L. D., Sultan, Florence, & Auld, F. The application of a psychometric scoring procedure to a group modification of the Thematic Apperception Test. J. consult. Psychol., 1955, 19, 83-89.

Franck, Kate, & Rosen, E. A projective test of masculinity-feminity. J. consult. Psychol., 1949, 13, 247-256.

Kagan, J. The measurement of overt aggression from fantasy. J. abnorm. soc. Psychol., 1956, 52, 390-393.

Lesser, G. S. Application of Guttman's scaling method to aggression fantasy in children. Educ. psychol. Measmt., 1958, 18, 543-552.

Lesser, G. S. Custom-making projective tests for research. J. proj. Tech., 1961, 25, 21-31.

McCandless, B. R., & Castaneda, A. Anxiety in children, school achievement, and intelligence. Child Develpm., 1956, 27, 385-90.

McClelland, D., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The achievement motive. New York: Appleton-Century Crofts, Inc., 1953.

Mandler, G., Sarason, S. B. A study of anxiety and learning. J. abnorm. soc. Psychol., 1952, 47, 166-73.

Mathias, R. Methodology and results of personality appraisal with Roschach, TAT and figure drawing. Paper read at American Psychological Association, Cincinnati, Ohio, September, 1959.

Michigan State Department of Mental Health. *Michigan Picture Test*. Chicago: Science Research Associates, 1953.

Murray, H. A. *Explorations in personality*. New York: Oxford University Press, 1938.

Murray, H. A. *Thematic Apperception Test Manual, and Thematic Apperception Test*. Cambridge, Massachusetts: Harvard University, 1943.

Mussen, P. H. *Handbook of research methods in child development*. New York: John Wiley and Sons, Inc., 1960.

Rabin, A. I. Devising projective methods for personality research. *J. proj. Tech*., 1961, 25, 6-10.

Rosenzweig, S., Fleming, Edith E., & Rosenzweig, Louise. The children's form of the Rosenzweig Picture-Frustration Study. *J. Psychol*., 1948, 26, 141-191.

Sarason, S. B. *Anxiety and learning*. *Human variability and learning*. Washington, D. C.: Association for Supervision and Curriculum Development, Dept. of National Education Assoc., 1961.

Sarason, S. B. A test anxiety scale for children. *Child Develpm*., 1958, 29, 105-113.

Sarason, S. B. *Anxiety in elementary school children*. New York: John Wiley and Sons, Inc., 1960.

Sarnoff, I. A cross cultural study of anxiety among American and English school children. *J. educ. Psychol*., 1958, 49, 129-36.

Taylor, Janet A. Drive theory and manifest anxiety. *Psychol. Bull*., 1956, 53-303-20.

Vernier, C., Whiting, J. F., & Meltzer, M. L. Differential prediction of specific behavior from three projective techniques. *J. consult. Psychol*., 1955, 19, 175-182.

# APPENDIX A

## Elementary Schools

This plan outlines a step by step procedure to direct the grouping in District A Elementary Schools.

1. Prior to the administration of the California Achievement Test, Battery (CAT) teachers will be requested to complete an official form in which they will recommend that each child in their class be placed in the regular developmental, or accelerated section for the coming year.[1] This form should be completed prior to testing so that a teacher's judgement will not be unduly influenced by the child's performance on the achievement test.

2. After the achievement tests have been scored, a roster for each sex and grade will be prepared listing the name of each child in order of his total score on the CAT. For example, the highest fourth grade boy would be placed first on the boy's list and so on until all children have been listed. Boys and girls will be placed on separate lists.

3. In order to set up the developmental section, the lowest ten boys and lowest ten girls on a basis of CAT total grade placement will be chosen. This will give an initial group of 20 children. If any children in this group are more than .5 grade placement points (one half year) above the rest of the group, these children will be placed in an average group. Every effort should be made to keep the ratio of boys to girls approximately equal. The procedure described above can be varied at the discretion of the principal but the group should not exceed 20.

4. The remaining children will be divided into the number of accelerated and average sections to be made up. For example, if the school has 120 fourth grade children, the lowest ten boys and the lowest ten girls would be placed in the developmental section and the remaining 100 would be divided into three sections of about 30 to 35 each.

5. To make up the initial accelerated section the top 15 boys and the top 15 girls will be tentatively chosen.

6. This will leave sufficient children to make up two average groups. These average groups would be chosen by placing the odd numbered boys and girls in one section and the even numbered in another. This procedure would result in the average children being heterogeniously grouped within the average range.

7. In schools having four or more sections in a given grade, two developmental sections or two accelerated sections may be formed if the achievement distribution, as compared with other schools in District A justifies this arrangement, with the approval of the Director of Elementary Education.

8. After this tentative grouping is completed, children in each group who are within .5 grade placement units of the nearest score of the next group will be checked and adjustments based on teacher recommendation will be made. For example, if a child is at the high boundary of the average group and the teacher recommendation is to place this child in the accelerated group, the child would be placed in the accelerated group. These adjustments would not be made except in cases of borderline children.

---

[1] School officials in District A used the terms Developmental, Regular, and Accelerated instead of slow, average, and superior employed in the Utah study.

9. In cases where the teacher's judgement disagrees radically with the test score (teacher recommends developmental placement, test recommends accelerated placement or vice versa), the scoring will be checked and if correct, the child will be retested with another form of the CAT.

   a. If the retest agrees with the first test, the child will be placed in accordance with the test results.
   b. If retest agrees with the teacher, the child will be placed in accordance with the teacher's opinion.
   c. If retest does not agree with either the initial test or teacher opinion, the child will be placed in an average group.

10. Records will be kept of children on whom the test and teacher disagreed radically and if a sufficient number of the cases occur, a special study of these pupils would be made.

11. Children transferring to District A schools without a satisfactory achievement score from other districts shall be tested with CAT Reading Test prior to placement in a class. Their total CAT grade placement should be corrected for the time of the year and they should be placed in an accelerated, average, or developmental section based on the corrected CAT grade placement total. If borderline on the CAT, the child will be placed in an average section.

   a. When the child reports to school, enrollment procedure will be completed and an appointment will be made for the child to take the placement test prior to class assignment.
   b. During the period between enrollment and class assignment the child will be placed in the orientation program.

## Secondary Schools

1. Grouping at the secondary school level will be carried out in classes in Mathematics and Science.
   a. The primary criterion for grouping in mathematics will be the student's score on the mathematics section of the California Achievement Test Battery.
   b. The primary criterion for grouping in science will be the combined score on the mathematics and language sections of the California Achievement Test Battery.
2. In order to maintain continuity between the curricula in the elementary and secondary school, the grouping procedure and the proportion of students in accelerated, regular, and developmental sections must remain comparable.
   a. Approximately 25% of the total number of sections in each course in mathematics and science should be designated as accelerated, 25% should be designated as developmental, and the remainder as regular classes. This percentage may be varied at the discretion of the principal if the distribution of scores indicates that more sections at the developmental or accelerated level are desireable.
   b. It should be remembered, however, that if the proportions of accelerated pupils at the Junior High School level are not comparable to the proportions at elementary level, the continuity of subject matter between the two levels may be disrupted. For example: If accelerated 6th grade pupils complete 7th grade mathematics and are then placed in a regular 7th grade math section when they enter Junior High School, these pupils will have to review material that they have already mastered.
3. Within each major grouping classification: accelerated, regular, and developmental, no attempt should be made to establish sub-groups on the basis of ability. In other words, all accelerated sections in 7th grade mathematics should be comparable in terms of ability. The make up of any particular section in this subject would be essentially random and would be determined primarily on the basis of scheduling efficiency.
4. An attempt should be made to keep the number of boys and girls in each section approximately equal.
5. The specific steps to be accomplished by the Junior High School or High School in determining the level of each pupil in mathematics and science are as follows:
   a. All boys at a given grade level will be listed in order of their score on the mathematics section of the California Achievement Test Battery.
   b. The top 25% will be identified for placement in accelerated mathematics sections.
   c. The lower 25% will be designated for placement in developmental sections and the remaining 50% for placement in regular sections.
   d. Girls will be designated for placement in accelerated, developmental, or regular sections in the same manner.
   e. The correct number of sections should then be scheduled so as to permit an enrollment of not more than 30 pupils in accelerated sections, not more than

35 pupils in regular sections, and not more than 20 pupils in developmental sections.

f. After determining the tentative placement of all boys and girls into accelerated, regular, or developmental sections, a number of pupils equaling 5% of the total number taking the class, whose scores fall closest to the borderlines between accelerated and regular sections and between regular and developmental sections will be rechecked against teachers' recommendations from the previous grade level.
   (1) In these cases, if the teacher's recommendation differs one placement level from the test recommendation the child will be placed in accordance with the teacher's recommendation.
   (2) If the teacher's recommendation differs two placement levels from the test recommendation, the child will be retested and his placement determined by the retest.

6. Students transferring from other school districts will be placed according to their achievement test scores as recorded on their transcript. If a satisfactory achievement test score is not available, the student should be tested using the California Achievement Test and placed according to the results.

## APPENDIX B

### Table D

### Means and S. D.'s on CMM and CAT Scores

| | | | | | | | First Battery<br>Weber County | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | CMM | | | | | | CAT | | | |
| | | | Language<br>Test 1 | | Non<br>Language<br>Test 2 | | Arith.<br>Test 3 | | Language<br>Test 4 | | Reading<br>Test 5 | |
| Tea. | Cd. | Grade | M | SD | M | SD | M | SD | M | SD | M | SD |
| 75 | (4th) | low | 14.9 | 6.55 | 34.0 | 9.40 | 3.81 | .45 | 3.99 | .66 | 3.35 | .65 |
| 81 | (4th) | low | 18.0 | 5.74 | 34.0 | 9.30 | 3.79 | .43 | 3.78 | .50 | 2.00 | .66 |
| 87 | (4th) | low | 9.4 | 4.53 | 35.0 | 10.67 | 3.83 | .47 | 3.72 | .48 | 3.27 | .56 |
| 76 | (4th) | average | 23.9 | 5.53 | 37.3 | 7.84 | 4.15 | .44 | 5.08 | .69 | 4.26 | .67 |
| 83 | (4th) | average | 26.4 | 6.89 | 42.1 | 8.14 | 4.55 | .43 | 4.95 | .61 | 3.43 | .68 |
| 88 | (4th) | average | 22.5 | 7.22 | 36.2 | 8.70 | 4.14 | .58 | 4.65 | .69 | 3.53 | .77 |
| 77 | (4th) | high | 31.0 | 9.61 | 40.8 | 6.97 | 4.78 | .47 | 5.56 | .81 | 4.86 | 1.00 |
| 82 | (4th) | high | 34.2 | 5.22 | 41.6 | 5.95 | 4.90 | .32 | 5.61 | .49 | 5.10 | .57 |
| 89 | (4th) | high | 32.4 | 7.41 | 41.7 | 6.84 | 4.92 | .55 | 5.82 | .82 | 5.04 | 1.07 |
| 78 | (6th) | low | 29.4 | 7.88 | 41.1 | 6.67 | 5.21 | .53 | 5.47 | .76 | 4.61 | .89 |
| 84 | (6th) | low | 27.7 | 8.23 | 42.9 | 7.61 | 5.10 | .66 | 5.13 | .81 | 5.38 | .88 |
| 90 | (6th) | low | 29.4 | 6.95 | 41.9 | 7.27 | 5.10 | .51 | 5.37 | .59 | 5.00 | 1.04 |
| 93 | (6th) | low | 28.2 | 5.65 | 39.3 | 8.24 | 5.00 | .65 | 5.11 | .87 | 4.47 | .88 |
| 80 | (6th) | average | 39.6 | 6.28 | 44.3 | 6.84 | 5.65 | .48 | 6.23 | .61 | 6.02 | .79 |
| 85 | (6th) | average | 40.3 | 7.09 | 44.3 | 7.66 | 6.06 | .45 | 6.35 | .79 | 5.38 | .88 |
| 91 | (6th) | average | 44.1 | 7.12 | 46.6 | 3.96 | 5.99 | .53 | 6.45 | .63 | 5.98 | .69 |
| 94 | (6th) | average | 38.4 | 6.89 | 43.5 | 6.34 | 5.70 | .46 | 6.06 | .63 | 5.86 | .82 |
| 79 | (6th) | high | 49.5 | 7.06 | 46.9 | 6.30 | 6.17 | .40 | 6.97 | .63 | 7.10 | .81 |
| 86 | (6th) | high | 51.4 | 5.67 | 50.1 | 5.04 | 7.14 | .59 | 6.99 | .39 | 6.95 | .69 |
| 92 | (6th) | high | 53.2 | 6.47 | 48.8 | 4.15 | 6.57 | .42 | 7.22 | .69 | 7.09 | .92 |
| 95 | (6th) | high | 48.3 | 6.26 | 48.8 | 6.03 | 6.32 | .43 | 7.02 | .70 | 6.88 | 1.02 |
| 60 | (7th) | low | | | | | 5.92 | .44 | 6.22 | .74 | 5.92 | .81 |
| 62 | (7th) | average | | | | | 6.56 | .92 | 7.53 | 1.46 | 7.13 | 1.28 |
| 61 | (7th) | high | | | | | 8.28 | .90/<br>.95 | 9.13 | 1.01 | 8.57 | .93 |
| 65 | (8th) | low | | | | | 5.88 | .55 | 6.05 | .96 | 5.89 | 1.21 |
| 66 | (8th) | low | | | | | 6.19 | .77 | 6.19 | .87 | 6.08 | .76 |
| 64 | (8th) | average | | | | | 7.57 | 1.20 | 8.04 | 1.00 | 7.81 | 1.39 |
| 68 | (8th) | average | | | | | 7.71 | 1.05 | 7.88 | 1.01 | 7.82 | .92 |
| 63 | (8th) | high | | | | | 9.92 | 1.10 | 10.16 | 1.02 | 9.45 | .81 |
| 67 | (8th) | high | | | | | 9.59 | 1.10 | 9.72 | .80 | 9.28 | .74 |
| 71 | (9th) | low Sci. | | | | | 7.07 | 1.11 | 7.18 | 1.01 | 7.26 | .94 |
| 74 | (9th) | low Math | | | | | 7.28 | .84 | 7.68 | .94 | 7.75 | .97 |
| 69 | (9th) | average Sci. | | | | | 8.31 | 1.19 | 8.93 | .92 | 8.54 | .98 |
| 72 | (9th) | average Math | | | | | 8.44 | 1.05 | 8.62 | .96 | 8.64 | 1.00 |
| 70 | (9th) | high Sci. | | | | | 10.48 | 1.09 | 9.96 | .72 | 9.74 | .55 |
| 73 | (9th) | high Math | | | | | 10.26 | 1.10 | 9.97 | .67 | 19.69 | .64 |

## Table H

### Means and S. D.'s on CMM and CAT Scores

First Battery

Ogden City

| | | | Test 1 | | Test 2 | | Test 3 | | Test 4 | | Test 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tea. | Cd. | Grade | M | SD | M | SD | M | SD | M | SD | M | SD |
| 23 | (4th) | | 25.2 | 7.31 | 39.3 | 6.65 | 4.30 | .48 | 4.57 | .84 | 3.99 | .88 |
| 28 | (4th) | | 25.7 | 7.59 | 34.5 | 8.90 | 4.47 | .54 | 4.93 | .78 | 4.44 | .90 |
| 29 | (4th) | | 23.0 | 8.29 | 39.1 | 7.02 | 4.46 | .55 | 4.85 | .95 | 4.01 | .89 |
| 30 | (4th) | | 29.8 | 8.16 | 38.1 | 6.99 | 4.76 | .48 | 5.05 | .98 | 4.53 | 1.02 |
| 34 | (4th) | | 24.9 | 6.39 | 37.0 | 8.54 | 4.30 | .42 | 4.91 | .93 | 3.95 | .89 |
| 35 | (4th) | | 26.1 | 10.68 | 32.1 | 9.48 | 4.53 | .55 | 5.05 | .82 | 4.65 | .99 |
| 38 | (4th) | | 26.9 | 8.25 | 39.5 | 8.12 | 4.66 | .48 | 5.26 | 1.02 | 4.89 | .87 |
| 41 | (4th) | | 27.2 | 9.70 | 40.4 | 8.63 | 4.69 | .50 | 5.04 | .85 | 4.38 | .77 |
| 42 | (4th) | | 27.3 | 8.61 | 39.3 | 7.53 | 4.45 | .51 | 4.76 | .70 | 4.32 | 1.03 |
| 43 | (4th) | | 28.1 | 7.12 | 39.2 | 6.96 | 4.41 | .43 | 4.49 | .77 | 3.84 | .72 |
| 48 | (4th) | | 26.5 | 8.26 | 39.9 | 7.31 | 4.54 | .55 | 5.11 | .90 | 4.61 | 1.06 |
| 49 | (4th) | | 28.6 | 9.51 | 36.8 | 7.81 | 4.75 | .52 | 5.08 | .87 | 4.67 | .90 |
| 50 | (4th) | | 28.7 | 8.45 | 39.4 | 5.57 | 4.74 | .57 | 5.21 | .88 | 4.36 | .99 |
| 26 | (6th) | | 38.6 | 9.58 | 43.0 | 6.89 | 5.90 | .49 | 6.24 | .80 | 5.88 | .84 |
| 31 | (6th) | | 42.7 | 10.35 | 46.6 | 5.97 | 5.84 | .71 | 6.42 | 1.13 | 6.23 | 1.38 |
| 32 | (6th) | | 42.6 | 8.77 | 46.2 | 5.93 | 5.91 | .57 | 6.10 | .85 | 5.77 | 1.11 |
| 33 | (6th) | | 39.1 | 11.05 | 47.7 | 6.29 | 6.09 | .67 | 6.16 | .96 | 6.11 | 1.28 |
| 36 | (6th) | | 37.4 | 9.60 | 44.5 | 6.63 | 5.75 | .52 | 5.92 | .86 | 5.92 | 1.19 |
| 37 | (6th) | | 36.3 | 9.72 | 44.4 | 5.89 | 5.65 | .65 | 5.88 | .86 | 5.68 | 1.36 |
| 39 | (6th) | | 42.8 | 10.36 | 46.1 | 6.11 | 6.04 | .85 | 6.40 | .90 | 6.13 | 1.37 |
| 40 | (6th) | | 39.4 | 10.68 | 47.1 | 6.69 | 5.99 | .63 | 6.13 | .85 | 5.81 | 1.01 |
| 44 | (6th) | | 44.3 | 9.61 | 45.8 | 6.27 | 6.24 | .81 | 7.10 | 1.12 | 6.48 | 1.32 |
| 45 | (6th) | | 47.0 | 8.87 | 47.8 | 4.23 | 6.26 | .58 | 6.75 | .66 | 6.78 | 1.21 |
| 46 | (6th) | | 44.5 | 10.05 | 46.3 | 6.31 | 5.96 | .63 | 6.99 | .87 | 6.61 | 1.37 |
| 47 | (6th) | | 45.9 | 8.81 | 48.3 | 4.87 | 6.41 | .42 | 6.72 | .78 | 7.03 | 1.09 |
| 51 | (6th) | | 40.1 | 12.36 | 43.7 | 9.71 | 5.88 | .82 | 6.15 | .79 | 5.88 | 1.46 |
| 54 | (6th) | | 45.6 | 8.00 | 46.7 | 6.57 | 6.22 | | 6.78 | .73 | 6.46 | 1.27 |
| 55 | (6th) | | 46.2 | 10.78 | 46.9 | 5.90 | 6.25 | .42 | 6.71 | .86 | 7.03 | 1.27 |
| 57 | (6th) | | 42.6 | 9.26 | 44.3 | 6.89 | 5.90 | .57 | 6.55 | .72 | 6.32 | 1.15 |
| 58 | (6th) | | 40.2 | 10.67 | 42.9 | 8.13 | 5.67 | .64 | 6.09 | .87 | 5.86 | 1.25 |

| | | | | | | | Arith. | | Lang. | | Reading | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Test 1 | | Test 2 | | Test 3 | | Test 4 | | Test 5 | |
| Tea. | Cd. | Grade | M | SD | M | SD | M | SD | M | SD | M | SD |
| 00 | (7th | Arith) | 41.0 | 6.58 | 29.8 | 4.87 | 7.48 | 1.17 | 8.12 | 1.06 | 8.71 | 1.04 |
| 01 | (7th | Arith) | 33.5 | | 29.8 | | 6.72 | | 6.98 | | 7.34 | |
| 02 | (7th | Arith) | 36.6 | 6.57 | 30.9 | 5.44 | 6.88 | .63 | 7.83 | .73 | 7.88 | .83 |
| 03 | (7th | Arith) | 26.4 | 9.18 | 24.1 | 5.86 | 6.05 | .84 | 6.32 | 1.33 | 6.77 | 1.40 |
| 04 | (7th | Arith) | 35.3 | 7.83 | 29.1 | 5.16 | 6.88 | 1.03/ .92 | 7.62 | 1.53 | 7.50 | 1.28 |
| 05 | (8th | Arith) | 47.9 | 9.41 | 33.8 | 5.62 | 9.20 | 1.24 | 9.33 | 1.06 | 9.57 | 1.20 |
| 06 | (8th | Arith) | 33.5 | 8.10 | 25.8 | 6.59 | 7.28 | 1.38 | 7.09 | 1.37 | 7.54 | 1.61 |
| 07 | (8th | Arith) | 39.7 | 5.33 | 29.9 | 6.25 | 8.26 | .71 | 8.85 | .78 | 9.34 | .87 |
| 08 | (8th | Arith) | 37.9 | 9.66 | 29.4 | 4.94 | 7.31 | 1.37 | 7.66 | 1.55 | 7.80 | 1.58 |
| 09 | (8th | Arith) | 46.0 | 8.84 | 31.1 | 5.58 | 8.89 | 1.18 | 9.40 | 1.11 | 9.54 | 1.22 |
| 10 | (8th | Arith) | 42.9 | 9.37 | 31.8 | 5.92 | 8.68 | 1.39 | 9.09 | 1.44 | 9.00 | 1.34 |
| 11 | (8th | Arith) | 27.6 | 9.75 | 26.5 | 5.65 | 6.38 | .91 | 6.76 | 1.63 | 6.54 | 1.48 |
| 12 | (8th | Arith) | 36.8 | 7.51 | 28.0 | 4.48 | 7.40 | 1.18 | 7.87 | 1.27 | 7.74 | 1.32 |
| 13 | (8th | Arith) | 37.4 | 8.15 | 30.1 | 6.24 | 7.43 | 1.10 | 7.81 | 1.68 | 8.11 | 1.56 |
| 14 | (8th | Arith) | 33.4 | 10.72 | 28.6 | 6.31 | 7.22 | 1.22 | 7.49 | 1.55 | 7.13 | 1.82 |
| 15 | (9th | Science) | 40.5 | 8.64 | 30.7 | 6.42 | 7.84 | 1.21 | 8.27 | 1.25 | 8.44 | 1.02 |
| 16 | (9th | Science) | 44.6 | 10.92 | 30.4 | 7.56 | 8.50 | 1.96 | 8.71 | 1.85 | 8.80 | 1.68 |
| 17 | (9th | Science) | 41.4 | 6.28 | 30.8 | 5.92 | 7.94 | 1.61 | 8.35 | 1.30 | 8.20 | 1.26 |
| 18 | (9th | Science) | 45.4 | 7.41 | 30.6 | 4.96 | 8.74 | 1.57 | 8.53 | 1.47 | 8.83 | 1.53 |
| 48 | (9th | Science) | 47.3 | | 34.4 | | 9.88 | | 9.17 | | 9.39 | |
| 19 | (9th | Algebra) | 45.4 | 9.70 | 33.9 | 5.60 | 8.80 | 1.47 | 8.69 | 1.14 | 8.96 | 1.06 |
| 20 | (9th | Algebra) | 38.4 | 8.42 | 28.9 | 5.02 | 7.99 | .99 | 8.60 | .96 | 8.32 | .92 |
| 21 | (9th | Algebra) | 40.4 | 8.69 | 32.4 | 8.03 | 8.32 | 1.43 | 8.68 | 1.30 | 8.49 | 1.18 |
| 22 | (9th | Algebra) | 47.3 | 9.50 | 33.6 | 5.87 | 10.00 | 1.29 | 9.86 | 1.16 | 9.61 | 1.08 |

Test 1: CMM Language
2: CMM Non-Language
3: CAT Arithmetic Total
4: CAT Language Total
5: CAT Reading Total

# APPENDIX C

## Hypotheses as Stated in Original Research Plan

Questions and/or hypotheses. All hypotheses have been stated in null form. Hypotheses will be tested for within-group, between-group, and sex differences.

4.1 There are no significant changes in achievement of superior, average, or below average students in the experimental and control systems during the course of the study.
4.2 There are no significant changes in the number of social isolates in groups of superior, average, or below average students during the course of the study.
4.3 The attitudes and opinions of teachers respecting ability grouping are not significantly related to the success of the ability-grouping program.
4.4 There are no significant changes in the number of under-achievers among superior, average, or below average students in the experimental and control systems during the course of the study.
4.5 There are no significant changes in attitude towards school among superior, average, and below average students in the experimental and control systems during the course of the study.
4.6 At the high school and junior high school level, students in ability-grouped classes in mathematics and science do not achieve more than they do in ungrouped subjects nor do they achieve more than similar students in ungrouped mathematics and science classes.
4.7 There are no significant changes in the number of behavior problems reported or the amount of truancy among superior, average, or below average students in the experimental and control systems during the course of the study.
4.8 There are no significant changes in school adjustment of superior, average, or below average students in the experimental and control systems during the course of the study.
4.9 There are no significant changes in the effectiveness of study methods among superior, average, and below average students in the secondary school experimental and control systems during the course of the study.
4.10 There are no significant changes in social adjustment scores of superior, average or below average students in the experimental and control systems during the course of the study.
4.11 There are no significant changes in self-perception with special reference to such variables as acceptance of self, inferiority feelings, level of aspiration, concept of self, etc., between superior, average, and below average students in the two systems.
4.12 There are no significant differences in the rate of achievement of superior, average, or below average students in the experimental system as compared with achievement of these same students prior to the establishment of ability grouping in the experimental system.
4.13 In random-grouped classrooms, inferior students have significantly fewer learning contacts, both in the classroom and out of the classroom with superior students than they have with other inferior students. Students generally tend to select both their friends and study companions from students with similar ability.

# APPENDIX D

# THE U.S.U. SCHOOL INVENTORY

| | |
|---|---|
| NAME | TEACHER'S NAME |

| | | |
|---|---|---|
| GRADE | SCHOOL | DATE |

BOY GIRL
(circle one)

## DIRECTIONS TO STUDENT

On the following pages you will find a list of questions concerning things about this school that you may or may not like. We should like to know what things about this school you like and what you dislike. Your answers will be secrets to yourself, and only those of us who are here giving the test will have a chance to see your answers. They will not affect your grade in any way. Your teachers will not see your answers. If you will answer these questions honestly and thoughtfully, the school will work to improve the conditions which your answers indicate need improvement.

There are no right or wrong answers. Indicate your answer by drawing a circle around "Yes", "No", or "?". Try to answer all questions either "Yes" or "No". If you are certain that you cannot answer "Yes" or "No", then use the question mark.

There is no time limit, but work rapidly.

Page 2

Yes No ? 1. Do you like all of the subjects you are now taking in this school?
Yes No ? 2. Does your teacher expect you to do too much work?
Yes No ? 3. Is it easy to make friends with your teacher?
Yes No ? 4. Does your teacher act as if she likes her work?
Yes No ? 5. Do you wish you could get better grades in school?
Yes No ? 6. Does your teacher tell you when you've done a good job?
Yes No ? 7. Does your teacher encourage you to do your best work?
Yes No ? 8. Do you have a hard time making new friends?
Yes No ? 9. Have you been able to get into the school activities that you like best?
Yes No ? 10. Have you found your teacher to be too strict?
Yes No ? 11. Would you like to study different things than the ones you are now studying?
Yes No ? 12. Are you proud of your teacher?
Yes No ? 13. Do you find it easy to make friends in this school?
Yes No ? 14. Does your teacher get upset over small matters?
Yes No ? 15. Are the things you are learning in school interesting?
Yes No ? 16. Do you feel you are popular with boys?
Yes No ? 17. Do you feel you are popular with girls?
Yes No ? 18. Can your teacher explain the lessons clearly?
Yes No ? 19. Do your principal and teacher act as if they are always right and you are always wrong?
Yes No ? 20. Are you frightened by the way your teacher calls on you in classes?
Yes No ? 21. Have you found it hard to prepare your lessons for your classes?
Yes No ? 22. Are you usually a leader in your group?
Yes No ? 23. Does your teacher require too much homework?
Yes No ? 24. If you were able to do so, would you like to attend some other school than the one you are now attending?
Yes No ? 25. Do you dislike your teacher?
Yes No ? 26. Is your principal too strict with students?
Yes No ? 27. Do you like your teacher's speaking voice?
Yes No ? 28. Do you feel that some of your teachers have held a "grudge" against you?
Yes No ? 29. Is there a small group of students who always plan class activities?
Yes No ? 30. Is your teacher lazy?
Yes No ? 31. Does your teacher have "pets" in this class?
Yes No ? 32. Is your teacher a good friend to the students?
Yes No ? 33. Do you have trouble keeping your mind on what you are studying?
Yes No ? 34. Are your lessons too long?
Yes No ? 35. Would you like to quit school now?
Yes No ? 36. Does your teacher play favorites?
Yes No ? 37. Are there any groups of children who won't let you play with them?
Yes No ? 38. Are you afraid of your teacher?
Yes No ? 39. Do the other students like you?
Yes No ? 40. Do you think the students in this school are "stuck-up"?
Yes No ? 41. Do you think that some of the women teachers in this school show favoritism toward boys in their classes?
Yes No ? 42. Is your teacher hard to understand?

Yes No ? 43. Are the students in this school fair in their play?
Yes No ? 44. Does school make you unhappy?
Yes No ? 45. Does your teacher make lesson assignments easy to understand?
Yes No ? 46. Are you smart in school?
Yes No ? 47. Is it easy to get to know your teacher?
Yes No ? 48. Are some of your lessons very boring to you?
Yes No ? 49. Does your teacher treat you fairly?
Yes No ? 50. Does your teacher admit it when she is wrong?
Yes No ? 51. Are the boys and girls in this school usually nice to you?
Yes No ? 52. Does your teacher give grades fairly?
Yes No ? 53. Is your teacher interested in you?
Yes No ? 54. Does your teacher show a lack of interest in class and school activities?
Yes No ? 55. Do you have difficulty keeping your mind on what goes on in class?
Yes No ? 56. Do your friends trust you?
Yes No ? 57. Does your teacher like to make you feel embarrassed before the class?
Yes No ? 58. Do you wish your teacher liked you better?
Yes No ? 59. Does your teacher really care whether you learn something in this class?
Yes No ? 60. Have you found that your teacher does not like to be with the boys and girls?
Yes No ? 61. Do you think that the boys and girls in this school like you as well as they should?
Yes No ? 62. Are you proud of your friends?
Yes No ? 63. Is your teacher often mean and unfair to you?
Yes No ? 64. Do your classmates usually feel that they know more than you?
Yes No ? 65. Do you feel that your teacher dislikes her job?
Yes No ? 66. Do you wish you could study better?
Yes No ? 67. Does your teacher treat you like you were a small child?
Yes No ? 68. Are you often left out of things other kids do?
Yes No ? 69. Do you think that this school is run as if it were a prison?
Yes No ? 70. Does your teacher understand you?
Yes No ? 71. Do your classmates seem to think you are not a good friend?
Yes No ? 72. Do you think your school requires too much homework?
Yes No ? 73. Is your teacher cheerful and pleasant?
Yes No ? 74. Do you like your lessons?
Yes No ? 75. Does your teacher give you enough individual help in your school work.
Yes No ? 76. Do you find it hard to be as popular as the other kids?
Yes No ? 77. Does your teacher lack a sense of humor?
Yes No ? 78. Do your classmates think you are smart?
Yes No ? 79. Is your teacher usually nice to you?
Yes No ? 80. Do you hate school?
Yes No ? 81. Is your teacher "bossy"?
Yes No ? 82. Do you prefer to be alone?
Yes No ? 83. Does your teacher like you?
Yes No ? 84. Do you find your school work dull and uninteresting?

Page 4
Yes No ? 85. Do you often think that your teacher does not like you?
Yes No ? 86. Are your feelings hurt easily?
Yes No ? 87. Do you wish you were smarter in school?
Yes No ? 88. Do your parents like your teacher?
Yes No ? 89. Is your teacher honest in her dealings with you?
Yes No ? 90. Do students who are not good at school work get treated fairly in this school?
Yes No ? 91. Do you often wish you had some other teacher?
Yes No ? 92. Do you worry about losing your friends?
Yes No ? 93. Is your teacher often so unkind or unfair that it makes you feel bad?
Yes No ? 94. Are there some students who do not like you?
Yes No ? 95. Are you lonesome in school?

## APPENDIX F

NAME____________________
GRADE____________________
SCHOOL____________________

## CLASSROOM ASPIRATION

The way you answer the following questions will tell how you rate yourself right now in a particular subject and how well you would like to rate in this particular subject.

CIRCLE ONE FOR EACH QUESTION

| | | | | | |
|---|---|---|---|---|---|
| 1a. How good are you at sports right now: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 1b. How good do you expect to become in sports: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 2a. How good are you at math right now: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 2b. How good do you expect to become in math: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 3a. How good are you at science right now: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 3b. How good do you expect to become in science: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 4a. How good are you at reading right now: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 4b. How good do you expect to become in reading: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 5a. How good are you at talking in front of the class: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 5b. How good do you expect to become at talking in front of the class: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 6a. How good are you as a leader of your classmates right now: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |
| 6b. How good do you expect to become as a leader of your classmates: | Nearly best in my grade | Pretty good | About average | Not so good | Nearly poorest in my grade |

## SOCIAL ASPIRATION

1. Look ahead and try to see yourself as you might be one year after you graduate from high school and answer the following:

A. How many social functions (parties, club meetings, dinners, etc.) do you expect to attend during an average month?

___________ more than 6
___________ 5-6
___________ 3-4
___________ 1-2
___________ none

B. How many times do you expect your name to appear in the newspaper in connection with a social affair during an average month?

___________ more than 6
___________ 5-6
___________ 3-4
___________ 1-2
___________ none

C. At what level do you expect to be an officer in an organization such as the Lions Club, Rotary Club, Country Club, Toastmaster's Club, or other similar organizations?

___________ President
___________ Vice President
___________ Minor Officer
___________ Just a member
___________ Not a member

2. Look ahead and try to see yourself as you might be five years after you graduate from high school and answer the following:

A. How many social functions (parties, club meetings, dinners, etc.) do you expect to attend during an average month?

___________ more than 6
___________ 5-6
___________ 3-4
___________ 1-2
___________ none

B. How many times do you expect your name to appear in the newspaper in connection with a social affair during an average month?

___________ more than 6
___________ 5-6
___________ 3-4
___________ 1-2
___________ none

C. At what level do you expect to be an officer in an organization such as the Lions Club, Rotary Club, Country Club, Toastmaster's Club, or some similar organization?

___________ President
___________ Vice President
___________ Minor officer
___________ Just a member
___________ Not a member

3. Please look at the descriptions below and answer the questions that follow. It is not necessary that every item under a description fit your family, just so most of the items generally describe your family.

A. This family earns more money than it can spend; college does not mean much to them; they belong to the"best" or most exclusive clubs; they often live in very large houses with large yards; they frequently have servants; they usually drive "big" cars (Cadillac, Lincoln, etc.).

B. This family has professional jobs (such as doctors and lawyers); they usually go to college and graduate; they are usually very active in clubs and organizations; they frequently live in large houses which they own; they frequently own two cars.

C. This family works for wages and salaries; they may own small businesses or farms; they usually spend most of the money they make; they sometimes use a college education as a means of getting them into social clubs and similar organizations; they frequently drive a relatively new car (not more than 4 or 5 years old); this is the typical "American family."

D. This family usually doesn't have much money; they work hard; the children usually do not go to college; they often live in rather small homes (3 or 4 rooms) which they may not own but rent; they often drive "older" cars (more than 4 or 5 years old.)

E. The job the father of this family has usually doesn't pay much money; the father frequently changes jobs; they usually drive a "real old" car (maybe 8 or more years old) or if they drive a newer car they have trouble paying for it; they are often "in hock".

3a. Check the above description which best fits your family now.

__________ A
__________ B
__________ C
__________ D
__________ E

3b. After you have married and have established a family and home, into which of the above categories do you think your family will fall?

__________ A
__________ B
__________ C
__________ D
__________ E

## EDUCATIONAL ASPIRATION

1. How important to you is getting a college education?

_____ Most important thing in my life
_____ Very important
_____ Quite important
_____ Not very important
_____ Not at all important

2. In terms of grades, what level are you trying to reach this year?

_____ In the top 10% of my class
_____ In the top one-third of my class
_____ A little better than average
_____ About average
_____ Below average

3. How far do you feel you must go in school to be satisfied with yourself?

_____ Continue in school beyond 4 years of college (like a professional school for doctors or lawyers)
_____ Graduate from college
_____ Attend college but not for the full 4 years
_____ Graduate from a trade school, business school, electronics school or some similar type school
_____ Graduate from high school

4. How many years of education do you expect to complete after high school?

_____ More than four years
_____ Four years
_____ Two years
_____ One year
_____ None

5. What level of education would you like your children to have?

_____ Go to school beyond 4 years of college (like a professional school for doctors or lawyers)
_____ Graduate from college
_____ Attend college but not for the full 4 years
_____ Graduate from a trade school, business school, electronics school or some similar type school
_____ Graduate from high school

6. What level of education do you expect your children to reach?

_____ Go to school beyond 4 years of college (like a professional school for doctors or lawyers)
_____ Graduate from college
_____ Attend college but not for the full 4 years
_____ Graduate from a trade school, business school, electronics school or some similar type school
_____ Graduate from high school

7. Are you planning to go on to school after high school?

_____ I am absolutely sure that I am
_____ Yes, definietly
_____ Probably
_____ Don't care
_____ No

8. Check the one description below that best tells your feelings about the expense of college.

_____ I would go to college even if I had to work all the way through
_____ I would go to college even if I had only enough money to get started and hope that something would happen so that I could continue
_____ I would not go to college unless I could be sure that I had enough money to get me through several years
_____ I would not start college unless I could be sure that I had enough money to get completely through college
_____ I would not go to college even if I had all the money that it takes

## APPENDIX G

Helps for scoring aggression (A)[a]

---

Opposite of aggression, submission, not to be scored

Words, expressions, and behavior indicating aggression

murdered or killed 5
hostile impulse
  carried out 4
  "felt like" 3
won't see the obvious 3
card rejected 3
inappropriate story 3
confused story 1
run away to punish parents
do purposeful wrong
always late, opposition
got a good talking to
can't think, not even going to try,
  and other indications of resistance
something that happened a while ago
unpleasant job assigned (qualify)
retribution
  he expects, D[a] (disappointment)
  he demands, A
wishful thinking
  only way to compete I[a]
  resistance A

mad 3
shoot
knock
hurt
hit
grab
push
beat
throw
steal

complain
gripe
scold
run
fight
swear
hate
lie
cheat

inadequate story
call names
break (violin)
blame others
run away
break
defy authority
violate rules
indifference
forget purposefully
better straighten up

I or A others blame him
I - A win every day
I need to win daily
I - A punishment for wrong doing
I can't do anything about it
I - A escape
I - A don't know how

dependent upon content
  fire
  sleep

Rejection of aggression, reconstruction, decreases score; reconstitution, insight, not simply a statement, but a real change in attitude; not denial, but more of a reform.
Denial of aggression--increases score.

---

[a] A--aggression, D--depression, I--inferiority feelings

Helps for scoring depression (D)$^a$

3 D's of depression: deprivation, denial disappointment

Reason is important. It may be I$^a$ of A$^a$.

suicide 5
disappointment 3
poverty of response 3
not happy, sad on 2 or 3, depending on reason
cry in disappointment
hopelessness
being left out
sleep A or D depending on reason
chronic inability to compete
duration, "has to do it a long time"
  insoluble task
parents not taking care of, neglect
lonely
something is always this way
die, deprivation
feels sorry for self
indifferent attitude

probably
I guess
maybe
could be
sad
tired
or something
being denied
  denying others, A
discouraged
worried
hurt
crying
nervous
upset
reactive depression
inability to deal with

To be depressed, you have to have a sense of values to be depressed about. It is safer to come up with aggression.

---

[a] A--aggression, D--depression, I--inferiority feelings

Helps for scoring inferiority feelings (I)[a]

---

- seeking parents' support
  - but not getting it
- need for success (unobtainable)
  - (if handled, OK)
- magic necessary for success (context)
- person has no control
  - (it finally happens)
- not able to compete
- unaccepted
- ridiculed: put on a dunce stool
- embarrassed (context)
- powerless to act
- scared
- lonely
- I A[a] wishful thinking
- A I escape
- A I run away, run
  - inadequacy, can't compete
- A I won every day
- A I never on time
- A I I don't know
- I A plans story so he is caught
- I gives up
- A blows up
- sick
- failed
- feels sorry
- confused
- boy identifying with girl
- trying
- cry
- ask question can't answer
- worry
- ashamed
- a dummy (person)
- staring
- kids laughing at him
- looking at him
- teasing
- complacency--dependent
  - upon context
- make fun of
- laugh at
- trying to play it (card 1)
- need for punishment
- masochism
  - to punish others, A
  - to gain acceptance, I
- day dreaming, wishing

---

[a] A--aggression, I--inferiority feelings

## Appendix E

Summary of IAV Means and Standard Deviations for Sample IV, Districts A and R Subgroups

| Group | Self Concept | | Ideal Self | | Discrepancy Score | | Acceptance of Self | |
|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD |
| District R | | | | | | | | |
| Superior boys | 45.60 | 3.85 | 50.00 | 2.39 | 4.60 | 3.27 | 50.10 | 6.76 |
| Superior girls | 45.80 | 3.72 | 49.80 | 1.78 | 4.10 | 3.33 | 48.70 | 8.41 |
| Average boys | 45.10 | 4.48 | 49.80 | 2.93 | 4.80 | 3.40 | 51.60 | 5.88 |
| Average girls | 45.10 | 4.46 | 49.70 | 1.66 | 4.90 | 3.93 | 50.40 | 7.82 |
| Slow boys | 45.20 | 4.52 | 49.50 | 3.51 | 4.80 | 3.93 | 50.40 | 9.45 |
| Slow girls | 45.30 | 4.19 | 49.00 | 2.73 | 4.00 | 3.64 | 49.80 | 8.54 |
| District A | | | | | | | | |
| Superior boys | 43.80 | 4.05 | 49.93 | 3.95 | 6.39 | 4.99 | 46.70 | 9.05 |
| Superior girls | 44.50 | 3.42 | 50.51 | 2.21 | 6.17 | 4.04 | 45.78 | 9.73 |
| Average boys | 44.54 | 5.20 | 49.63 | 3.58 | 5.72 | 4.53 | 47.44 | 9.08 |
| Average girls | 44.12 | 4.04 | 49.57 | 1.96 | 5.77 | 3.68 | 46.84 | 8.04 |
| Slow boys | 44.53 | 4.58 | 48.96 | 3.58 | 4.66 | 3.82 | 47.66 | 9.55 |
| Slow girls | 44.39 | 5.82 | 49.62 | 2.65 | 6.11 | 5.52 | 46.21 | 10.89 |